This e-book is published in the CAA series *Computer Applications and Quantitative Methods in Archaeology*.
Human Computer Interaction, Multimedia, Museums

15 Towards Collaborative Decipherment of Non-Verbal Markings in Archaeology
   Barbara Rita Barricelli, Stefano Valtolina, Giovanna Bagnasco Gianni and Alessandra Gobbi

21 Archaeological Documentation in the Field: the Case of the Roman Forum of Cástulo
   Ana Martínez Carrillo, Marcelo Castro, Francisco Arias de Haro and Manuel Serrano

30 Implications for the Design of Novel Technologies for Archaeological Fieldwork
   Tom Frankland and Graeme Earl

37 OpenArcheoSurvey, or ‘Being Educated by the Digital Fieldwork Assistant’
   Jitte Waagen, Nils de Reus and Rogier Kalkers

48 The Use of iPad as a Documenting Tool on an Archaeological Excavation on Govče 2011 Project in North - Eastern Slovenia
   Eva Butina

57 Back into Pleistocene Waters. The Narrative Museum of Casal de’ Pazzi (Rome)
   Augusto Palombini, Patrizia Gioia, Antonia Arnoldus-Huyzendveld, Marco Di Ioia and Sofia Pescarin

66 Etruscanning 3D: an Innovative Project about Etruscans
   Eva Pietroni, Daniel Pletinckx, Wim Hupperetz and Claudio Rufo

77 Personalizing Interactive Digital Storytelling in Archaeological Museums: the CHESS Project
   Laia Pujol-Tost, Maria Roussou, Olivier Balet and Stavroula Poulou

91 Installation for Interpretation of Archaeological Sites. The Portus Visualisation Project
   Javier Pereda

102 Material Motion: Motion Analysis for Virtual Heritage Reconstruction
   Kirk Woolford and Stuart Dunn

110 Interactive Workspace for Exploring Heterogeneous Data
   Uros Damnjanovic and Sorin Hermon

Simulating the Past

120 The Use of CFD to Understand Thermal Environments Inside Roman Baths: A Transdisciplinary Approach
   Taylor Oetelaar, Clifton Johnston, David Wood, Lisa Hughes and John Humphrey
Structural Assessment of Ancient Building Components, the Temple of Artemis at Corfu
Georg Herdt, Aykut Erkal, Dina D’Ayala and Mark Wilson Jones

Final Results of the Virtual 3D Reconstruction of the East Pediment of the Temple of Zeus at Olympia
András Patay-Horváth

Teaching Cultural Heritage and 3D Modelling through a Virtual Reconstruction of a Medieval Charterhouse
Andres Bustillo, Ines Miguel, Lena Saladina Iglesias and Ana Maria Peña

3D Reconstruction in Archaeological Analysis of Medieval Settlements
Daniele Ferdani and Giovanna Bianchi

Handling Transparency in 3D Reconstructed Online Environments: Aquae Patavinae VR Case Study
Daniele Ferdani, Bruno Fanini, Guido Lucci Baldassari, Ivana Cerato, Sofia Pescarin

3D Documentation for the Assessment of Underwater Archaeological Remains
Barbara Davidde Petriaggi, Roberto Petriaggi, Gabriele Gomes de Ayala

Post-Excavation Analysis in Archaeology Using 3D-Technology: the Case Study of Hala Sultan Tekke
Kostas Anastasiades, Sorin Hermon, Nicola Amico and Giancarlo Iannone, Karin Nys

A New Approach for Interactive Procedural Modelling in Cultural Heritage
René Zmugg, Ulrich Krispel, Wolfgang Thaller, Sven Havemann, Martin Pszeida and Dieter W. Fellner

Virtual Reality Simulations in Cultural Heritage
Ioanneta Vergi

Taking Excavation to a Virtual World: Importing Archaeological Spatial Data to Second Life and OpenSim
Isto Huvila and Kari Uotila

Using ConML to Visualize the Main Historical Monuments of Crete
Panagiotis Parthenios

A High-Performance Computing Simulation of an Irrigation Management System: The Hohokam Water Management Simulation II
John T. Murphy

Field and Lab Recording

Application of RTI in Museum Conservation
Eleni Kotoula

Automatically Recognizing the Legends of Ancient Roman Republican Coins
Albert Kavelar, Sebastian Zambanini and Martin Kampel
<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>Multispectral Imaging of Historic Handwritings</td>
<td>Fabian Hollaus</td>
</tr>
<tr>
<td>258</td>
<td>Multispectral Image Analysis of a Censored Postcard from 1942</td>
<td>Florian Kleber, Fabian Hollaus and Robert Sablatnig</td>
</tr>
<tr>
<td>264</td>
<td>Semantic Web Technologies Applied to Numismatic Collections</td>
<td>Ethan Gruber, Sebastian Heath, Andrew Meadows, Daniel Pett, Karsten Tolle and David Wigg-Wolf</td>
</tr>
<tr>
<td>275</td>
<td>Automatic Coin Classification and Identification</td>
<td>Reinhold Huber-Mörk</td>
</tr>
<tr>
<td>295</td>
<td>Intra-Site Analysis and Photogrammetry: the Case Study of the ‘Buca Di Spaccasasso’ (Grosseto, Italy) an Eneolithic Funerary Site</td>
<td>Giovanna Pizziolo, Daniele Pirisino, Carlo Tessaro and Nicoletta Volante</td>
</tr>
<tr>
<td>308</td>
<td>Site Recording Using Automatic Image Based Three Dimensional Reconstruction Techniques</td>
<td>Victor Ferreira, Luís Mateus and José Aguiar</td>
</tr>
<tr>
<td>316</td>
<td>Photographic Rectification and Photogrammetric Methodology Applied to the Study of Construction Process of Provincial Forum of Tarraco</td>
<td>M. Serena Vinci</td>
</tr>
<tr>
<td>324</td>
<td>Image-Based 3D Documentation of Archaeological Trenches Considering Spatial, Temporal and Semantic Aspects</td>
<td>Robert Wulff and Reinhard Koch</td>
</tr>
<tr>
<td>337</td>
<td>Digital Photogrammetry: a Contribution to the Study of Early Middle Ages Sarcophagi Quarries of Panzoult (Indre-et-Loire, France)</td>
<td>Daniel Morleghem</td>
</tr>
<tr>
<td>344</td>
<td>Low-Cost Photogrammetry and 3D Scanning: the Documentation of Palaeolithic Parietal Art in El Niño Cave</td>
<td>Alejandro García Moreno and Diego Garate</td>
</tr>
<tr>
<td>350</td>
<td>3D Documentation in Archaeology: Recording Las Cuevas Site, Chiquibul Reserve, Belize</td>
<td>Fabrizio Galeazzi, Holley Moyes and Mark Aldenderfer</td>
</tr>
<tr>
<td>363</td>
<td>Social Spreading of Geometric, Recorded Data from a Range of Types of 3D Scanners via a Web Data Server</td>
<td>Jorge Angas and Paula Uribe</td>
</tr>
</tbody>
</table>
386 3D Model of the Roman Walls of Lugo (Galicia, Spain) Using a Terrestrial Laser Scanner and an Unmanned Aerial Vehicle
   María J. Iniesto-Alba, Alicia Canizares-Sánchez, David Miranda and Rafael Crecente

398 (Re)seeing the Engraved Block of El Mirón Cave (Ramales de la Victoria, Cantabria, Spain)
   Vera Moitinho de Almeida, Luis Teira, Manuel González-Morales, Lawrence G. Straus, Millán Mozota and Ana Blasco

406 Meshlab as a Complete Open Tool for the Integration of Photos and Colour with High-Resolution 3D Geometry Data
   Marco Callieri, Guido Ranzuglia, Matteo Dellepiane, Paolo Cignoni and Roberto Scopigno

417 Enhancing Surface Features with the Radiance Scaling Meshlab Plugin
   Xavier Granier, Romain Vergne, Romain Pachon, Pascal Barla and Patrick Reuter

422 OpeninfRA – Storing and Retrieving Information in a Heterogeneous Documentation System
   Alexander Schulze, Frank Henze, Felix F. Schäfer, Philipp Gerth and Frank Schwarzbach

432 Towards Reverse Engineering Archaeological Artefacts
   Vera Moitinho de Almeida and Juan Anton Barceló

Data Analysis, Modelling and Sharing

444 ARCA: Creating and Integrating Archaeological Databases
   Maria del Carmen Moreno Escobar

457 A Database for Radiocarbon Dates. Some Methodological and Theoretical Issues about its Implementation
   Igor Bogdanović, Juan Antonio Barceló and Giacomo Capuzzo

468 Standardised Vocabulary in Archaeological Databases
   Matthias Lang, Geoff Carver and Stefan Printz

474 Modelling Imperfect Time in Datasets
   Koen Van Daele

480 Distribution Analysis of Bone Remains in the Prehistoric Site of Mondeval De Sora (Belluno - Italy): Issues and Proposals
   Maria Chiara Turrini, Federica Fontana, Antonio Guerreschi and Ursula Thun Hohenstein

487 Places, People, Events and Stuff; Building Blocks for Archaeological Information Systems
   Paul J. Cripps

498 ArcheoInf, the CIDOC-CRM and STELLAR: Workflow, Bottlenecks, and Where do we Go from Here?
   Geoff Carver

509 @OccupyWatlingStreet: Can we find out Who was occupying What, Where and When in the Past?
   Keith May
Connecting Archaeology and Architecture in Europeana: the Iberian Digital Collections
Ana Martínez Carrillo, Arturo Ruiz and Alberto Sánchez

Open Access Journals in Archaeology and OpenAccessArchaeology.org
Doug Rocks-Macqueen

SVG Pottery: Upgrading Pottery Publications to the Web Age
Stefano Costa

Through an Urban Archaeological Data Model Handling Data Imperfection
Asma Zoghlami, Cyril de Runz, Dominique Pargny, Eric Desjardin and Herman Akdag

Guerrilla Foursquare: a Digital Archaeological Appropriation of Commercial Location-Based Social Networking
Andrew Dufton and Stuart Eve

Conceptualising eScience for Archaeology with Digital Infrastructures and Socio-Technical Dynamics
Teija Oikarinen and Helena Karasti

Geospatial Technologies and Analysis

Intrasite Spatial Analysis of the Cemeteries with Dispersed Cremation Burials
Marge Konsa

A Specific Approach for a Peculiar Site: New Spatial Technologies for Recording and Analysing a Palaeolithic Site (the Cave of La Garma, Northern Spain)
Alfredo Maximiano, Pablo Arias and Roberto Ontañón

Use of Quantitative Methods to Study an Alpine Rock Art Site: the Mont Bego Region
Thomas Huet

“The Whole is More than the Sum of its Parts”- Geospatial Data Integration, Visualisation and Analysis at the Roman Site of Ammaia (Marvão, Portugal)
Eleftheria Paliou and Cristina Corsi

Scattered Chronology - Surface Artefact Survey and Spatial Analysis of Ceramic Concentrations
Ondrej Malina and Jakub Silhavy

Ecological and Social Space in the High Mountains in South Norway 8500 – 2000 BP
Espen Uleberg and Ellen Anne Pedersen

Chalcolithic Territorial Patterns in Central Moldavia (Iaşi County, Romania)
Robin Brigand, Andrei Asăndulesei, Olivier Weller and Vasile Cotuigă

Settlement Patterns in Drahany Uplands (Czech Republic): GIS and Quantitative Methods Based Approach
Lukáš Holata
Rural Life in Protohistoric Italy: Using Integrated Spatial Data to Explore Protohistoric Settlement in the Sibaritide
Kayt Armstrong and Martijn van Leusen

Reconstructing the Ancient Cultural Landscape Around Pompeii in 2D and 3D: from Scientific Data to a Computer Animated Museum Exhibit
Sebastian Vogel, David Strebel, Michael Märker and Florian Seiler

Using GIS to Reconstruct the Roman Centuriated Landscape in the Low Padua Plain (Italy)
Michele Matteazzi

Integrating Spatial Analyses into Foraging Societies Land Use Strategies. A Case Study from the Nalón River Basin (Asturias, North of Spain)
Alejandro García, Miguel Angel Fano and Diego Garate

Lost Worlds: A Predictive Model to Locate Submerged Archaeological Sites in SE Alaska, USA
Kelly R. Monteleone, E. James Dixon and Andrew D. Wickert

Familiar Road, Unfamiliar Ground. Archaeological Predictive Modelling in Hungary
Gergely Padányi-Gulyás, Máté Stibrányi, Gábor Mesterházy and Mártin Deák

Mathematical Models for the Determination of Archaeological Potential
Nevio Dubbini and Gabriele Gattiglia

Calculating Accessibility
Irmela Herzog

Simulated Paths, Real Paths? A Case Study of Iberian Cessetania (Iron Age Society)
Joan Canela Gràcia

Open Source GIS for Archaeological Data: Two Case Studies from British and Egyptian Archaeology
Anna Kathrin Hodgkinson, Luca Bianconi and Stefano Costa

Speeding up Georeferencing with Subpixel Accuracy
Gianluca Cantoro

Multi+ or Manifold Geophysical Prospection?
Apostolos Sarris

Managing Data from Multiple Sensors in an Interdisciplinary Research Cruise
Øyvind Ødegård, Martin Ludvigsen, Geir Johnsen, Asgeir J. Sørensen, Stefan Ekehaug and Fredrik Dukan and Mark Moline

Towards Detection of Archaeological Objects in High-Resolution Remotely Sensed Images: the Silvretta Case Study
Karsten Lambers and Igor Zingman
ArcheOS and UAV P, a Free and Open Source Platform for Remote Sensing: the Case Study of Monte S. Martino ai Campi of Riva del Garda (Italy)
Alessandro Bezzi, Luca Bezzi, Rupert Gietl and Nicoletta Pisu

The Visualization of the Archaeological Information through Web Servers: from Data Records on the Ground to Web Publication by Means of Web Map Services (WMS)
Julio Zancajo, Teresa Mostaza and Mercedes Farjas

Theoretical Approaches and Context of Archaeological Computing

Crafting Archaeological Methodologies: Suggesting Situational Method Engineering for the Humanities and Social Sciences
César Gonzalez-Perez and Charlotte Hug

Boundary Concepts For Studying the Built Environment. A Framework of Socio-Spatial Reasoning for Identifying and Operationalising Comparative Analytical Units in GIS
Benjamin Vis

Everything Flows: Computational Approaches To Fluid Landscapes
Dimitrij Mlekuž

Reliability of the Representation of a Distribution: a Case Study on Middle Bronze Age Metal Finds in the Seine Valley
Estelle Gauthier and Maréva Gabillot

Assessing Positional Uncertainty due to Polygon-to-Point Collapse in the Cartographic Modelling of Archaeological Scatters
Fernando Sanchez and Antoni Canals

Theoretical Space-Time Modelling of the Diffusion of Raw Materials and Manufactured Objects
Estelle Gauthier, Olivier Weller, Jessica Giraud, Robin Brigand, in collaboration with: Pierre Pétrequin and Maréva Gabillot

A Tangible Chronology
Jean-Yves Blaise and Iwona Dudek

Reconstructing Fragments: Shape Grammars and Archaeological Research
Myrsini Mamoli and Terry Knight

Grammar Modelling and the Visualisation of an Uncertain Past: the Case of Building 5 at Portus
Matthew Harrison, Simon Keay and Graeme Earl

Can Infovis Tools Support the Analysis of Spatio-Temporal Diffusion Patterns in Historic Architecture?
Jean-Yves Blaise and Iwona Dudek
History in 3D: New Virtualization Techniques for Innovative Architectural and Archaeological Scholarship and Education
James C. Sweet, Krupali Krusche, Christopher R. Sweet, and Paul Turner

Investigating the Effectiveness of Problem-Based Learning in 3D Virtual Worlds. A Preliminary Report on the Digital Hadrian’s Villa Project
Lee Taylor-Nelms, Lynne A. Kvapil, John Fillwalk and Bernard Frischer

Building Blocks of the Lost Past: Game Engines and Inaccessible Archaeological Sites
Anna Maria Kotarba-Morley, Joe Sarsfield, Joe Hastings, John Bradshaw and Peter Nicholas Fiske

Re-reading the British Memorial: A Collaborative Documentation Project
Nicole Beale and Gareth Beale
Foreword

This volume is an extension of the printed volume “Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012. It consists of a selection of the peer-reviewed papers presented at the Computer Applications and Quantitative Methods in Archaeology 2012 conference hosted by the Archaeological Computing Research Group at the University of Southampton, UK between 26th and 30th March 2012. The conference included 53 sessions divided between the themes of simulating the past, spatial analysis, data modelling and sharing, data analysis, management, integration and visualisation, geospatial technologies, field and lab recording, theoretical approaches and the context of archaeological computing, and a general theme. In addition there were 12 workshops. A total of 380 papers and posters were presented, and two key note addresses. Alongside the lively conference atmosphere at the venue there was a thriving social media back channel. In addition to these proceedings there is therefore a broad ranging multimedia record of the event, accessible via the conference website.

The co-organisers of CAA2012 and myself would like to thank the CAA Steering Committee for their advice and assistance. We are also indebted to Professor Anne Curry (Dean of Faculty of Humanities) and Professor Jonathan Adams (Head of Archaeology) for their support and encouragement. Many individuals and organisations in Southampton and further afield, including the sponsors and exhibitors, contributed to making the conference such a success. Of course without the many delegates travelling from across the globe and offering such exciting contributions there could have been no conference, and we are very grateful to them for their lively contributions to all aspects of the event. Finally, we would like to offer our thanks to the superb team of volunteers that made CAA2012 possible. The Archaeological Computing Research Group at Southampton was very proud indeed to be able to host the 40th CAA conference and we know that this was demonstrated by the enthusiasm, dedication and professionalism of the postgraduate and undergraduate students that gave so much of their time to the event.

I very much hope that you enjoy these proceedings and all the many related outputs from CAA2012, and I look forward to seeing you at future CAA conferences.

Graeme Earl

Southampton, United Kingdom, November 2012
Human Computer Interaction, Multimedia, Museums
Towards Collaborative Decipherment of Non-Verbal Markings in Archaeology

Barbara Rita Barricelli, Stefano Valtolina, Giovanna Bagnasco Gianni and Alessandra Gobbi
Università degli Studi di Milano, Italy

Abstract:
This paper presents an approach to collaborative decipherment of non-verbal markings in archaeology. The approach aims at enabling collaboration among archaeologists through knowledge management and annotation tools. The implementation of the approach and of multimedia information retrieval strategies in the IESP (International Etruscan Sigla Project) System is presented. The system was designed and developed in the frame of an International archaeological project focused on the study of Etruscan sigla (non-verbal markings) found on objects discovered in different digging sites distributed in the Mediterranean area.

Keywords:
Cultural Heritage, Knowledge Management, Ontology-Based Model, Collaboration

1. Non-verbal Markings

Unlike what happens in deciphering verbal languages, in the case of non-verbal signs it is possible to study their elements from a graphical point of view and to apply similarity techniques to support the human interpretation activity. The root of our research, that involves computer scientists and archaeologists, is the participation to an international project (IES Project) that involved American and Italian Universities. In particular, the archaeological environment considered by the project was the Etruscan one. In what follows we will use the Etruscan case to explain the motivations of our choices in the definition of our approach. However, the existence of non-verbal markings in other ancient languages makes the approach applicable to other fields of research, with the only requirement of the adaptation of the terminology to the more suitable one.

As to Etruscan language, thousands of examples of non-verbal markings exist. Typically they are referred to as graffiti, a term that is found to be inadequate. Instead, the Latin word siglum (pl. sigla) – corresponding to the Greek one sema (pl. semata) – should be used. Examples of Etruscan sigla are given in Figure 1. Etruscan sigla, composed by one or more symbols, numbers or letters, are dated from around 700 BCE to the first century BCE. They are incised, painted or stamped on different types of objects; e.g., pottery weights, spindle whorls, sarcophagi, burial urns, roof tiles, architectural terracotta, boundary stones, stone walls, and a wide variety of artefacts in bronze (axes, fibulas, helmets, knives, razors, sickles). The contexts in which the objects have been found include cemeteries, sanctuaries, ports, artisans' quarters and habitations – all spheres of Etruscan life and afterlife. The digging sites are distributed in the Mediterranean area: In Italy, from the heartland of Etruria, to Etruscan expansion areas on the Bay of Naples and near Bologna and the Po Valley; but also in Greece, Crete and Turkey. The study of Etruscan sigla is aimed at assessing the real consistency of archaeological indicators according to a deductive method that takes into account a dialectic comparison between the ideas of function and role (de Grummond et al. 2000; Bagnasco et al. 2009; Gobbi in press). The function of an object could be in fact be deduced by its shape. On the other hand, the role of the same kind of object can be determined differently on the basis of the conditions of their discovery and from the comparison of iconographic sources. This means that the meaning of sigla can change widely according to the context in which they have been discovered.

Corresponding author: barricelli@di.unimi.it
An example is the case of V-shape siglum that can be interpreted as a number 5 or letter U. The same uncertainty exists in interpreting a siglum formed by a cross inscribed in a circle: it could mean the Greek letter theta or could be the graphic representation of a sacred space (Bagnasco 2008).

The experience we developed in the frame of IES Project led to the design and development of an approach and its software implementation for:

1. Analyzing cases of recurrent sigla as cultural indicators of non-verbal communication within their different archaeological contexts.

2. Supporting questions about function and role in the field of sigla and according to a multifaceted perspective that takes into account archaeological data to a larger extent.

The main goal of the approach and the final system is to assess sigla with reference to their geographical range and chronology, to the nature of the objects and contexts to which they belong and to the layout of the graphic design. The enormous amount of data, the variety of the cultural background of archaeological experts involved the wide span of different hypotheses about the interpretation of each siglum type and their relationships led to the design of a tool that supports collaborative activities and dialectic comparisons. With 'collaborative' we mean that the involved people decide to work together to reach a shared goal. In our case the goal is to interpret the meaning of non-verbal markings by means of the comparison of images, the sharing of descriptions, and the collaborative contribution by whole archaeologists’ community.

2. Archaeology and Technology

The understanding of past civilizations and their influence on the modern world and life passes inexorably through the study of their systems of signs among which is language. Several attempts to automatic deciphering of lost languages have been made. Knight and Yamada (Knight and Yamada 1999) developed a computational approach for unknown scripts decipherment. Their approach is based on a study of phonetic and written scripts in verbal languages. Another approach that tried to implement automatic decipherment is the one presented by Snyder et al. (2010). They defined a statistical model that incorporates a range of linguistics intuitions in a statistical framework. The most recent contribution in this research area is the one given by Rao et al. (2009). The authors used an algorithm and its implementation in a software system to measure the entropy in Indus script. The results of this measurements were compared with the entropy of other types of languages – i.e. natural
Towards Collaborative Decipherment of Non-Verbal Markings in Archaeology
Barbara Rita Barricelli et al.

(e.g. Tamil, English), artificial (FORTRAN programming language), and non-linguistic (e.g. human DNA). The comparison among the results led the authors to the conclusion that Indus is probably a language but do not help in decipher the script.

Our approach, instead, is not aimed at finding ways of automatic decipherment but, as will be presented in the rest of the paper, at supporting the archaeologists in unravelling systems of signs implemented in different cultural environments that only seldom might have a wider dissemination. Moreover, as in the case of the above mentioned representation of the sacred space (the cross inscribed in a circle) such a wider dissemination might be only attached to only one sign. Therefore, our approach reflects more what was only theoretically proposed in (Puyol-Gruart 1999). The author presents the main challenges that arise from archaeological research in ancient languages decipherment and some ideas to address them by applying computer science and artificial intelligence. In particular, Puyol-Gruart suggested the use of Knowledge Discovery in Databases (KDD), Visual Information Management (VIM) and Multi-agent Systems (MAS). Moreover our research focuses on the study of non-verbal markings that is incised, painted or stamped symbols, numbers or letters found on different type of objects discovered in several cultural contexts that can be considered essentially non-literary. Therefore, our approach is not devoted to decipher linguistic signs but to trigger analysis processes for interpreting the meaning of non-verbal markings by means of comparisons of images, descriptions and collaborative contributions of the whole archaeologists’ community. Another important contribution of our work is the collaborative nature of our approach and its implementation. The active participation of all the archaeologists involved in a research group, as well as the one who work in other groups, is in fact very important for a successful management of a shared knowledge base and for markings decipherment.

3. Collaborative Knowledge Management in Archaeology

The Cultural Heritage (CH) domain and especially modern archaeology would significantly benefit from the access to shared knowledge bases in order to better support distributed and asynchronous collaboration. Several problems arise from the integration of different archives, and one of the most important is the need of establishing a common knowledge representation to be used to exchange data among all the stakeholders involved in the collaboration. To this aim, the approach presented in this paper (Valtolina 2008) uses CIDOC-CRM (Croft et al. 2005) to describe the specific archaeological domains (e.g., Etruscan, Roman, and Egyptian) and to retrieve the associated context information from distributed data sources. Being CIDOC-CRM suitable to each CH context, in our approach this model is used as backbone for defining a situated ontology (Floyd and Ukena 2005) able to fit all cultural aspects of the given archaeological domain. To offer a common representation of a specific knowledge base, in our approach the ontology schema is also used to integrate information of heterogeneous data sources. The elements defined in each logical schema are expressed according to the classes of the ontology and to a proper set of mapping information. A detailed explanation about the integration strategy adopted in our approach is reported in (Valtolina 2008; Aliaga et al. 2011). Concepts related to non-verbal markings have been integrated in the ontology schema: 1) The new sigla concepts – e.g., siglum context information, siglum typology, siglum position, direction and orientation, siglum images and descriptions and information about the possible combinations of sigla; 2) Support – i.e. the object on which the non-verbal marks are placed; 3) Component – i.e. the part of the object (e.g., neck, foot, and side); 4) Issue – i.e. the non-verbal marks that can be divided in sigla, decoration, and textual inscription. The two last concepts are important because, although they are not proper sigla, they have
been often found in association with sigla and so they have to be considered together as parts of a whole system. The use of the ontology provides a universal point of access to all the information distributed in the available data sources. A semantic mediator permits to express queries in terms of domain’s concepts rather than using entities defined in databases’ logical schemas – e.g., retrieve all sigla and relate them with findings or monuments stored in diggings or museum archives. The ontology unifies scattered archives containing sigla and diggings information, and gives the archaeologists the possibility to formulate new hypotheses about their research and to re-build a context (according the function and role of the findings on which the sigla have been discovered). As an example of new hypotheses formulation, if a set of similar sigla have been discovered in a specific place and all of them have the same chronology then it is possible to deduce that this siglum is a brand and it is possible to trace its geographic distribution in order to understand something about its circulation. An example of context re-building is the one in which a siglum is similar to others belonging to findings discovered in sacred places (i.e. tombs) and then it is possible to infer that this siglum has a religious meaning even if the origin of the related finding is unknown.

4. IESP System

IESP system has been designed by applying a participatory design approach. It implements collaborative knowledge management by offering annotation tools to be used for exchanging opinions about the classification of sigla. The system allows to apply multimedia research strategies based on similarity checks between different pictures and drawings of sigla. The conditions of similarity used by the system are based on recurrent patterns of marking that constitute the outcome of several years of research carried out by different groups of archaeologists. The features implemented in IESP System, and their combined use, allow archaeologists to complex systemic readings of the sigla stored in different databases that may lead to new insights and lines of investigation. IESP System allows to search sigla on the base of textual descriptions, estimated date, provenance and relationships with other sigla.

IESP System allows also to search sigla on the base of textual descriptions, estimated date, provenance and relationships with other sigla. It integrates Oracle Text (Shea et al. 2007) full-text retrieval strategies, a technology included in the Oracle11g DBMS. IESP System also addresses the research based on a thematic dictionary, i.e., a thesaurus, that contains terms used for describing the shape of siglum (e.g. alphabetiform, numeriform, symbol) or other features such as typology (e.g., craticula, tridens acutus, forma quadrans) or technique (e.g., incised, painted, stamped). Moreover, the thesaurus is used to put in relation each of its terms with their synonyms or to define hierarchical or semantic relationships among them. These relationships are used to create a semantic network of terms organized according to different conceptual associations. Examples of these associations are:

• SYN (synonyms): hook SYN hanger.
• BT (broader term): sacred place BT tomb.
• NT (narrower term): tomb NT sacred place.
In this way, the system is able to retrieve descriptions that contain relevant texts by expanding the queries to include similar or related terms as defined in the thesaurus. Another feature implemented in IESP System is the image retrieval and object recognition tool that is aimed at classifying image content. Such feature process the information contained in the pictures and drawings of a siglum and create an abstraction of its content expressed in terms of visual attributes. These visual attributes concern textures, primitive shapes distributions (i.e., the segmentations of the images in simple geometrical shapes) and the information about placements of shapes and textures in the image. This image retrieval feature is developed using MultiMedia services (Pelski et al. 2007) that integrate the storage, retrieval, and management of media files in Oracle11g DBMS.

5. Evaluation and Future Work

For evaluating the prototype, we organized a two-day workshop in Tarquinia. The aim was to involve all the stakeholders in a collaborative activity of evaluation and re-design of the prototype and benefit from the insights provided by them all. The number of participants to the workshop was ten. The workshop activity was divided into five main parts: (1) We asked the participants to fill in a preliminary questionnaire aiming at collecting information about participant’s background knowledge and everyday use of Information Technologies. The questionnaire also asked about which aspects of technology (e.g. functionality, easiness of use, aesthetics, etc.) were considered to be more relevant by the users. (2) A user test during which the participants were invited to use the system for the classification of three sigla consecutively and were asked to report their impressions and difficulties by thinking aloud during the interaction. During the tests, we captured screenshots and audio of the sessions. (3) A focus group that aimed at providing all the participants with a common understanding and a shared terminology about the process of classification of sigla, both when it is done manually and when it is supported by technology. (4) A second focus group aiming at collecting impressions about the functionalities already implemented in the system and suggestions for its re-design. During this activity, we collected information both through a moderated discussion and through an exercise in which each participant was asked to write on different post-it notes the five most relevant benefits and the five most important limitations of the system. Post-it notes were then clustered on the basis of the topic of their content and allowed to define a number of aspects on which future design cycles should focus on. (5) A meeting during which the definition of the thesaurus implemented in the system was discussed. Workshop participants reported that functionality, easiness of use and the possibility of sharing contents with colleagues and with the scientific community are key assets for a system built for supporting the work of an archaeologist. The data collected during the user test and the focus groups showed a high level of consistency and allowed us to identify two main areas of intervention for the re-design of the system: a) the procedure of sigla classification, with the purpose of offering a more suitable set of attributes, and b) the design of the interface, in order to add new functionalities and make the system more appealing. Further studies are expected to focus on innovative information retrieval solutions for supporting semantic researches combining heterogeneous data sources in order to put in relation information of sigla with information about the discovery contexts of the supports on which the sigla have been found. The idea is to design a semantic engine using which the archaeologist can submit semantic queries for information that is not contained in, or cannot be searched in only one database but that has to
be recovered combining knowledge contained in the network of the integrated databases using the ontology after a period of testing held by a community of archaeologists.

Acknowledgments

This work was supported by the ITN “Marie Curie Actions” entitled “DESIRE: Creative Design for Innovation in Science and Technology” (reference: PITN-GA-2008-215446-DESIRE). The authors acknowledge Nancy de Grummond and thank all the people who participated to the workshop.

References


Archaeological Documentation in the Field: the Case of the Roman Forum of Cástulo

Ana Martínez Carrillo
Universidad de Jaén, Spain

Marcelo Castro, Francisco Arias de Haro and Manuel Serrano
Conjunto Arqueológico de Cástulo, Spain

Abstract:
The aims of this contribution are the approach of a methodology for recording archaeological information as part of the excavation work on the forum located at the ibero-roman city of Cástulo (Linares, Jaén) and the presentation of some preliminary results. First, we describe the spatial and temporal context of the archaeological site. Next, we present the recording methodology used during the archaeological field work. Then the initial results are sketched, and finally conclusions and proposals for the future investigation of the site are set out.

Keywords:
Field Documentation, Novel Technologies

1. Introduction

As Sprague (1982) pointed out; an archaeological intervention generates several types of documentation that can be classified into field work documentation, analytical documentation, administrative documentation (reports) and scientific publications. From these, the documentation of fieldwork must be highlighted as the basis for building the others. Fieldwork documentation usually includes sketches of the investigated area, maps and plans at different scales, inventories of materials, structures forms to describe their relationships, forms to describe the stratigraphy, photographs, videos, etc. These documents contain contextual and relational data needed to construct the historical sequence and to test scientific hypotheses, and are therefore of vital importance in archaeological research.

Archaeological records and their interpretation are the basis of the archaeological discipline. Records not only include documentation of all materials recovered in an archaeological context, but also contain their textual and graphical information in order to reconstruct the three-dimensional context with all its aspects. Documenting a particular context properly through graphical and textual information is essential for the development of scientific hypotheses and to explain and to give meaning to a historic space.

During the past 10 years, a range of prototype, bespoke and/or commercial applications for the digital documentation of excavation have been developed in different countries and on the basis of distinct fieldwork practices (Kvamme 1999, 164-167; Lock 2003, 78-123). Also, Clarke et al. (2003) and Powlesland et al. (1998) provide some interesting examples on information flow through to publication).

Although the functionality that these provide is gradually improving, a more active role for such digital methods in the actual process of archaeological interpretation is still pursued. Digital datasets can encourage reflexivity in the excavation process, as they enable easier correlation, re-assessment and re-assembly of fragmented information in the excavation archive (Hodder 1999, 178-188).
2. Temporal and Spatial Context: the Ibero-Roman City of Cástulo and the Project Forum MMX

The Iberian oppidum of Cástulo is one of the largest fortified settlements of the Iberian Peninsula dating from Antiquity. Its extension has been estimated at 50 ha. This settlement is located on a smooth plain facing south, situated on the right bank of the Guadalimar river. In addition to the area located inside the wall, the archaeological site includes an extensive suburban environment estimated at more than 1.800 ha. where cemeteries, roads, a port and workshops from the iberian-roman period have been found, in addition to other settlements belonging to the Prehistoric and the Middle Ages.

The importance of the archaeological site of Cástulo (Linares, Jaén) lies in its origins and in the continuity of occupation during thousands of years. This is due mainly to its strategic geographical location that gives it clear visual control of the Guadalquivir valley and the north area. The intense occupation through the centuries, and the various archaeological interventions carried out in recent years give us a great deal of knowledge of the successive phases of occupation.

However, the ibero-roman period is the best known, not only for the importance acquired during this time, but also for being the period most documented in the archaeological interventions. From this period the work of Professor Blázquez Martínez must be highlighted in the sanctuary of La Muela (Blázquez and Molina 1975a), dated around s. VIII BC and the iberian and roman cemeteries (Arribas and Molina 1969; Molina and Blázquez 1973; Blázquez and Molina 1975b; Remensal and Blázquez 1979; Canto 1979, Canto and Urruela 1979).

The relevance of this site is unquestionable and it is essential to highlight the prominence which Cástulo deserves, both locally and regionally. Forum MMX project was developed along these lines, and its scientific objectives focus on the historical reconstruction of an episode key in shaping the current European identity, as was the constitution of the Roman Empire, which in its day meant the cultural and political consolidation of a large region around the Mediterranean. The recovery of this project through the forum of the ancient ibero-roman city of Cástulo will also impact the value of this archaeological heritage.

At the socio-political level, this project includes the collaboration of many researchers who come to perform studies of materials and plan archaeological excavation activities and, therefore, to be personally responsible for the development of research. At the same time a large segment of the citizens of Linares have been” particularly active as “cultural volunteers.
in recent years in the realization of these tasks. Therefore, the reconstruction through research and over time of the physical space of the forum will also involve the recovery of this patrimony for the citizens of the 21st century, and in fact give extraordinary value to the archaeological monuments (a public area of 70 ha, located 6 km from Linares, Jaén).

The scale of the stated scientific objective and the heterogeneity of the research team mean that we required the ad hoc development of a proper system of registration and documentation to ensure the quality of the information collected and provide further scientific exploitation. To this end we resorted to the use of new technologies and the computerization of the archaeological record in its different phases (fieldwork, laboratory analysis and publication). In this contribution the initial interest is in the first phase of archaeological work: the field record.

3. Computerization of the Fieldwork. State of the Art

The practical nature of fieldwork often means doing something else at the same time as recording data. Usually, this activity is intimately connected with the observation or measurement itself.

Archaeological observation proves to be very complex to model in database terms as it presents a broad variety of analytical objects, concepts and actions that are related in a wide variety of ways (Madsen, 2003). During the research process both relations and object definitions require re-adjustments as new conceptual categories (e.g. a broader stratigraphic group or phasing) emerge as part of the interpretive process. Furthermore, typological constraints and uncertainty about basic material properties (e.g. colour or chronology), caused either by the differences in excavation recording methods or simply the subjective nature of archaeological description, make the task of defining archaeological units and their characteristics even more complex.

For this reasons, the implementation of an excavation data model within a georelational data framework requires advanced linkages between textual and graphical information (D’Andrea 2003).

Fieldworkers need to spend as much time as possible in observation, and have only a limited attention capacity to deal with recording, be it on a paper form, a tape recorder, a camera or a handheld computer.

In this sense, several applications have been designed for recording all the observations during the archaeological fieldwork. As example, the “FieldNote system” was designed to support the full range of fieldwork-related activities from initial planning through data collection to collation, display and analysis (Ryan and Pascoe 1999).

One step forwards in this kind of applications are those which develop 3D models from GPS positions taken at the top and the bottom of the excavation units boundaries. The 3D geometric modelling consists of representing geometric and spatial relationships of volumetric objects. It could be helpful in the context of archaeological excavation units representation and analysis.

Once excavation units are geometrically modelled, it is possible to refer them within a trench or the entire archaeological site, to handle them in various ways (zoom, rotation, translation), to perform on them 3D spatial analysis such as volumetric calculus or intersection computation, to make various kinds of queries such as to find out excavation units that have a certain number of artifacts, to generate sections anywhere in the 3D model, and finally to publish it. As well as improving data analysis techniques, if this 3D modelling operation can be done during the excavation, it could greatly help archaeologists to plan more
efficiently their daily excavation strategy. This methodology has been successfully developed on the archaeological site of Tell ‘Acharneh (Syria) (Losier et al. 2007).

Also the methodology used in the excavation of Paliambela Kolindros (Greece) shows a powerful and wholly effective tool for 3D intra-site documentation and analysis. This systematic exploration of archaeological reasoning processes within the domains of ontology, visualization and time can reveal new opportunities for the organization, analysis and representation of archaeological data in GIS environments (Katsianis et al. 2008).

All this works show not only the importance of the recording of the spatial features in the archaeological fieldwork, but also the creation of a computer system for unifying all the contextual information (textual and graphical).

The evolution of computing is playing an important role here, since in the field phase it can accelerate the registration and processing of data and at the same time the information collected can be applied to the overall strategy of the excavation. For the subsequent laboratory work, the new methodology allows easy control of the recorded data and a qualitative improvement of the general entries. On the other hand, an improvement in the analysis of the data would make the relationship between various data sets easier. In short, computerization of archaeological records preserves a certain dynamic in the excavation files and ensures their conservation at the level of document archives.

4. Methodology

4.1 Registration of the Textual Information and its Associated Graphical Information

In the case of the Forum MMX project, a field inventory system called Too Waste has been designed. The main objective of this system for recording information in the field is to systematize and standardize the information that can be obtained from an archaeological intervention. The importance of standardization of the information lies in the possibility of being reinterpreted retrospectively, allowing data to be accessed by different researchers and enabling the development of new hypotheses about the same historical space.

This system allows the capture of textual and attached graphical information during the fieldwork. For this task different forms have been designed for recording aspects such as the type of deposit, the documented materials and the excavation process.

The first unit defined is the volumes. A volume is defined by coordinates (X, Y) to which levels are associated with vertical (Z), which are those formed by strata in their different sequential levels. This formula distinguishes the different documented volumes as superficial level, division by construction, complete or arbitrary division of space. To each of the registered levels it is possible to associate an image of the same. The X,Y,Z data capture has been carried out using a Total Station Theodolite.

The surfaces that make up the volumes can also be described in this system. The information collected from them are the volumes and levels that are associated with vertical (Z), which are those formed by strata in their different sequential levels. This formula distinguishes the different documented volumes as superficial level, division by construction, complete or arbitrary division of space. To each of the registered levels it is possible to associate an image of the same. The X,Y,Z data capture has been carried out using a Total Station Theodolite.

The surfaces that make up the volumes can also be described in this system. The information collected from them are the volumes and levels that are associated with vertical (Z), which are those formed by strata in their different sequential levels. This formula distinguishes the different documented volumes as superficial level, division by construction, complete or arbitrary division of space. To each of the registered levels it is possible to associate an image of the same. The X,Y,Z data capture has been carried out using a Total Station Theodolite.

The surfaces that make up the volumes can also be described in this system. The information collected from them are the volumes and levels that are associated with vertical (Z), which are those formed by strata in their different sequential levels. This formula distinguishes the different documented volumes as superficial level, division by construction, complete or arbitrary division of space. To each of the registered levels it is possible to associate an image of the same. The X,Y,Z data capture has been carried out using a Total Station Theodolite.

Also a small description field and the possibility of associating a video have been added.

Finally, the records were divided into three-dimensional records or individual records. Three-dimensional records are spatially referenced on the volume that
contains them. This form specifies the type of content and treatment of materials, being able to associate a picture of the process, detail and/or end.

The form for individual records contains the same information as the three-dimensional records, but with the difference that in these registers its exact position has been marked in order to be able to reproduce it later, hence we assigned coordinates X, Y, Z.

The computerization of the data collected in these forms has been carried out as follows: Data are collected on paper forms which are completed on Black Pad using a digital pen and sent via internet to be interpreted by OCR and stored in a database where the data are available for consultation, edition and export for use in external applications.

The general scheme of the system is given in Fig. 2. The system components are:

- Paper forms.
- Digital pen: Linked via Bluetooth to a mobile phone (Android or BlackBerry).
- Black Pad Router: Installed on the mobile terminal. It sends the forms to the processing platform.
- Black Pad server: Responsible for validating and processing the shipment, generating an OCR file from its contents and sending the result to the central server.

Once each data set is received at the central server Too Waste identifies the type of form and the page number, extracting the various items covered in the type of form. If it is a new shipment the corresponding code is generated from a counter and shown to the user on the terminal screen.

In addition to the processing of the data written on the forms, on some of them it is possible to attach images to items. These images are taken using the camera’s own Android or BlackBerry device and attached to send data. The attached images are also processed and added to each shipment so they are available for viewing by the user in the Internet control panel.

Therefore, for this data collection work in the field the replacement of traditional inventory sheets of paper by electronic tokens stored on a mobile device has been tested. Through the use of computerized data on a mobile device the majority of the information is codified, thus streamlining data collection in the field. The stored information goes directly into a database accessible on line, allowing the correction of erroneous records and the combination of data.
from different teams working simultaneously in the excavation area.

4.2. The Recording of Graphical Information

While information technology has had a major impact on the analysis and dissemination of data from archaeological interventions (database of archaeological sites, GIS analysis, implementation of web pages, blogs ...), it has had less impact on the phase of data collection in the field. However, there have been some efforts to acquire graphical information using digital methods such as photogrammetry (Barceló et al. 2002), correction of photographs (Reali and Zoppi, 2001), the pairing of pictures (Avern, 2001a), conducting archaeological drawings from given points for a season total (Schaich, 2002), the use of 3D modelling (Avern, 2001b) and 3D laser scanning of an archaeological area (Doneus and Neubauer, 2004).

However, despite great experimentation with different techniques, none of these methods has been used conventionally, as none has proven to be a complete solution in terms of speed, simplicity, accuracy and accessibility. In all these previous experiences it should be emphasized that the latest trends in the capture of graphical information in the field are characterized by the generation of 3D models. Within this line of action there are currently available techniques which are low cost, flexible and capable of generating 3D models (Pollyfays et al. 2003). This allows rapid and complete records to be generated in circumstances unfortunately common in archaeological interventions such as extreme environmental conditions and little time and money (Avern and Franssens 2011).

In the proposed Forum MMX a charting system has been tested based on the rapid acquisition of data and the low cost of devices used. The registration of graphical information has followed the following process:

- **Data Capture**

  The data capture method is fast and simple; for each of the areas excavated it is enough to take several perimetal photographs of the area being excavated. The same procedure is repeated each time the excavation level is lowered.

  The greater the number of pictures taken the more information the 3D model will have, but we must also bear in mind that it will thus generate a larger file.

- **Information processing**

  The pictures are then processed with ARC 3D software, the images sent to a server and the 3D model generated from them. The process can take minutes or hours, depending on the size of the photographs taken and sent. This software also allows to preview the generated 3D model (Vergauwen and Van Gool, 2006).

- **Edition of the 3D model**

  Finally the generated model was edited using the free software Meshlab, which is a free and open source that can process and edit 3D unstructured triangular meshes. The system is designed to facilitate the task of 3D data processing, which usually consists of cleaning, inspection and visualization of 3D meshes. The generated model allows 3D visualization of horizontal and vertical sequences of a given area.
5. Conclusions

First of note is the importance of documentation and preservation of information during the registration work in the field. This is essential because of the destructive nature of archaeological excavation and therefore the ephemeral nature of the information.

As a result of this the permanent nature such documentation should have is undeniable. This information is the basis for assumptions and approaches which will define and form a particular site and therefore should be as objective and homogeneous as possible, to be used as often as is necessary.

At this point a critical reflection of the inner workings of the archaeological discipline with respect to the methodology of recording the information in the field becomes necessary. Following on from this idea the Too Waste registration system has been designed, in which information from an archaeological intervention is made simpler and more objective. The field log data contain valuable information, being the usual way of capturing field notes, drawings and photographs which researchers perform during a procedure. Through Too Waste we can obtain a highly accurate graphical description of the components that form the archaeological context (volumes, surfaces, layers of materials and records). This detailed recording enables further 3D virtual reconstruction, since the exact positions of the elements have already been taken. Another advantage of this system is its easy accessibility. All data can be readily accessed by researchers over the Internet. The system based on client/server structure has many functions, such as workflow customization, document query and management, workflow supervision and function control.

On the other hand, graphic documentation has been generated which allows the realization of a 3D model from photographs taken in a determined area and in turn the realization of 2D graphic documentation (scale drawings of plants and stratigraphic profiles) from a 3D model. This represents a quantum leap in the quality of graphic information preparation, as the usual method is the realization of 3D reconstructions from excavation drawings. The chosen system is also inexpensive in economic terms, so it can be extended to the great majority of archaeological operations. In short, the development and consolidation of this system aims at developing a tool for use in future proceedings in the archaeological zone of Cástulo, with the longer-term aim of achieving long-term consistency of the documentation obtained in other archaeological operations.

References


Implications for the Design of Novel Technologies for Archaeological Fieldwork

Tom Frankland and Graeme Earl
University of Southampton, UK

Abstract:
This paper describes some of the first results from a research project which aims to develop new technologies to enhance how fieldwork archaeologists communicate. Based on the results of several field studies, interviews and a technology probe, a number of implications for designing novel technologies to support archaeological fieldwork are discussed. The stimulus for this research comes from recognising that the primary motivation for developing new technologies for archaeological fieldwork is to either acquire or manage data. While this is perhaps unsurprising, given that archaeological research is a destructive process, new and original technologies have the potential to improve how archaeological research is conducted in a multitude of ways. This research therefore examines how novel technologies might benefit archaeological field practice by exploring two alternative themes - how technology can support communication and enhance the awareness of archaeologists working in the field.

Keywords:
Fieldwork, Technology, Design, Communication, Awareness, Human-Computer Interaction

1. Introduction

This paper introduces some early results from research conducted as part of the RCUK DE PATINA project (http://www.patina.ac.uk). The overall aim of the PATINA project is to design technologies for supporting research. Our contribution to the project, and the subject of this paper, examines this aim in the context of a single domain, archaeological fieldwork, and explores how archaeological research practices could benefit from the design of novel technologies. While reviewing the literature published within the discipline that relates to the development or introduction of novel technologies for fieldwork, it is apparent that in the large majority of cases technology has been introduced with the primary purpose of supporting the acquisition or management of archaeological data. Given that archaeological research is a destructive process and that the primary research material cannot be re-examined, this is not considered surprising. However, the potential for technologies to support archaeological research extends far beyond these applications.

One area we are interested in supporting through technology is communication. Communication is vital in an archaeological fieldwork setting and plays a central role in archaeological interpretation. Formal conversation, informal discussions at the tea break or in the pub, and even idle chat with the archaeologist digging alongside you all help to shape an interpretation of the evidence (May and Crosby 2004). Technologies introduced to support the acquisition or management of archaeological data may help archaeologists to do their job more efficiently, but they may also reduce the need for archaeologists to communicate with one another. In contrast, we recognise that technologies should promote the significant role conversation plays in the interpretive process.

Communication, and the flow of information around a site, is strongly influenced by the social hierarchy. Archaeological fieldwork teams are often structured in a very hierarchical manner, which can form a social divide between the archaeologists who produce the data through excavation and those who interpret it (Berggren and Hodder 2003). Archaeologists with roles primarily associated
with excavating often feel disillusioned, perceiving their work as little more than unskilled manual labour (Everill 2008). We hope that a technological intervention might expose this social hierarchy by directing attention towards it, influencing both archaeologists’ communication and behaviour. We are also interested in whether technology can be used to empower archaeologists lower down the site hierarchy, allowing for alternative, egalitarian interpretations of the archaeological evidence, in addition to the traditional forms of interpretation currently produced.

Our approach for researching these themes draws on disciplines such as human-computer interaction, where the development of novel technologies is a primary research goal. The methods we chose to explore archaeological fieldwork practice - a combination of field studies, interviews and ‘probes’ - are methods frequently deployed in HCI in order to build up an understanding of user requirements and stimulate ideas about the design space.

The primary research methodology employed throughout this research was field study, which took an ethnographic approach to exploring how archaeologists communicate on site, their awareness of one another’s activities, and their work practices which exist around both existing and novel technologies. There have been numerous previous ethnographic studies conducted of archaeological excavations, notably the work of Edgeworth (2006). In contrast to these ethnographies, the field studies described here used methods deployed frequently by human-computer interaction practitioners. These methodologies, particularly ethnomethodology (Garfinkel 1967), were considered to be more appropriate given the aim of the ethnography was to aid the design process. Furthermore, the findings from the field studies are presented here as a series of implications for design (Dourish 2006), illustrating how the results of the studies directly inform the design process.

As archaeologists we often assume we have a good knowledge of the behaviour and work practices of others in the field. However, there are many reasons why this is not the case. The research aims of an excavation, the material being excavated, and the personalities and experiences of the team assembled mean that no two excavations can be regarded as the same. Furthermore, as archaeological practitioners, there is a good chance that we no longer consciously perceive the mundane activities we engage in on a day-to-day basis, and when asked to recall these, would probably struggle to do so. Ethnographers have long recognised the importance of viewing a familiar situation as an ‘outsider’, and the increase in valuable insights this can yield. Finally, conducting ethnography allows us to make a motivated choice of what to study, viewing an environment through a particular ‘lens’. For example, a HCI practitioner might choose to pay particular attention to users’ interaction with their environment, or, as in this case, their social interactions with one another.

2. Field Studies

2.1 Portus

The first field study was conducted at the research excavations at the Roman port of Portus, at Fiumicino near Rome (http://www.portusproject.org). Portus is an important archaeological site that the University of Southampton has been surveying and excavating for over a decade. The field study was conducted over a period of five days. Staff and students from the University of Southampton were interviewed and work practices at the site were observed and recorded.

One of the most surprising findings from the field study was the observation that conversations about archaeology were more or less likely to occur depending upon what activities the archaeologists were currently engaged in. It was noticeable that when two
or more archaeologists were actively engaged in some task while situated together in the same trench, they were unlikely to discuss the archaeology or features of that area. However, if one of these archaeologists stood up, or moved out of the trench to stand on the edge, conversations about the archaeology were much more likely to occur. One hypothesis might be that having a higher vantage point allows archaeologists to see ‘the bigger picture’, and facilitates the observation of subtle patterns in the archaeology. Alternatively this may function metaphorically – with the distance encouraging reflection, without necessary introducing any practical benefits to visual interpretation.

A similar scenario was also seen to occur when an excavation supervisor would approach a trench where other archaeologists were working. On the majority of these occasions, the supervisor would remain on the edge of the trench, discussing the archaeology with the archaeologists working within. If one of these archaeologists became increasingly engaged in this conversation, they were often observed to stand up or move to join the supervisor on the edge of the trench as well. In addition to this, once a conversation was established between two archaeologists on the edge of the trench, other archaeologists would move to join them. These conversations were witnessed to grow quite large, and would occasionally involve five or six members of the excavation team. Numerous questions abound from this behaviour. For example, does this movement in some way act as a gesture to others, inviting them to join the archaeologist in reflecting on the archaeology? Or might this movement stimulate the archaeologists to think in a particular way? Alternatively there may simply be a perception among archaeologists that a conversation held on the edge of the trench is more ‘archaeological’ and therefore more attractive.

Another behaviour that was interesting to observe at Portus was the amount of time spent by members of the excavation team in areas of the site other than the one they were allocated to work in. This time was often spent either in conversation with other members of staff or examining the contents of the various trenches across the site. Based on the informal interviews conducted at the time, it appeared that this behaviour was highly beneficial to the archaeologists, as it not only improved their personal knowledge of the site but also their ability to interpret how the area they had been allocated to work in related to the emerging picture of the site as a whole.

An additional field study was conducted at Portus by fellow PATINA researcher Angeliki Chrysanthi. This study was based around deploying small head-mounted cameras known as ‘Looxcie’ to archaeologists as they carried out everyday tasks while at the excavation. The Looxcie not only records footage continuously but also offers the user the ability to clip the last 30 seconds of video to a file. This functionality was considered appealing as it allowed the archaeologists to ‘bookmark’ moments of significance to them. Several of the videos captured over the duration of the trial featured an archaeologist explaining the latest interpretation of the excavation area to their colleague wearing the Looxcie, demonstrating that these conversations were of enough significance for an archaeologist to consider ‘archiving’ them. In the interviews that followed the study, some of the archaeologists raised concerns about capturing video at an excavation. In particular, they were uncomfortable wearing the device, as they felt they were treated differently by their colleagues while filming them. A publication discussing these findings and more is currently in preparation by Angeliki Chrysanthi.

2.2 Pompeii

Following the success of the Portus field study, a smaller field study was conducted at Pompeii based around a team of archaeologists from the Pompeii Archaeological Research Project: Porta Stabia (PARP:PS) and from the
Pompeii Quadriporticus project (PQP). The main reason for conducting this study was to assess how the team of archaeologists at Pompeii has benefitted from replacing traditional paper-based methods of field recording with Apple’s iPad. Staff and students were again interviewed and observed while working on site.

The majority of archaeologists at Pompeii, including the site director, suggested in their interviews that the introduction of the iPad has positively benefitted personal communication across the site. However, the majority of archaeologists that were interviewed also struggled to pinpoint specific reasons for why this had occurred, while the director referred to this phenomenon as an ‘intangible’ improvement to communication. One obvious suggestion is that there is a level of hype associated with using the iPads, although this might have been expected to diminish after three years of using them. Another possibility is that the iPad provides archaeologists with the ability to build conversations around an immediate, shared point of reference. This might not encourage more interaction, but the ability to view associated data, information or visuals might allow richer and longer conversations to develop. This was seen to occur in particular among the architectural survey team in their early morning meetings.

2.3 Itchen Abbas

The most recent field study to be conducted was at a student field school at Itchen Abbas, near Winchester. For many of the students who attended this field school, it was the first time that they had experienced any practical archaeology. Staff and students were again interviewed and observed over several days.

On the first day of the study, the students were taken on a site tour provided by the supervisors and director. During this short tour, the students were informed of the site’s historical significance, the archaeology of several test pits was explained and they were shown good and bad technique with tools and equipment. Following the tour, most of the students interviewed felt that they had a good understanding of the overall archaeology of the site. Speaking to students the next day yielded a very different response. Few students had any idea of what was emerging in trenches other than the one they were working in. Those who did had discussed their trench with their friends working in different trenches at the tea and lunch breaks. It emerged that the reason the other students were unaware of developments elsewhere on site was a fear of moving from their trench, in case the supervisor saw them and scolded them for not working. This provides an interesting contrast to the discoveries made during the field study at Portus, where the staff are more experienced and the aims of the excavation significantly differ.

In addition to the field studies and interviews, a technology probe was deployed at the field school. A technology probe (Hutchinson et al. 2003) is essentially a technological deployment designed to collect data that inspires new design; they are not prototypes to be iterated, but aim to provoke interesting responses from the users (Boehner et al. 2007). The technology probe deployed at Itchen Abbas was conducted by placing small spherical markers wherever archaeologists stood by the edge of a trench for a predetermined amount of time. As the markers accumulated,
it was hoped that the archaeologists would become increasingly aware of their own and others behaviour around the trench (Fig. 1). This is the concept Garfinkel (1964) refers to as a ‘breaching experiment’. The use of spherical physical markers was a deliberate one, as it was thought that these would best parallel how locations are marked in ‘augmented reality’ apps. In this way, the deployment of the technology probe might be considered an ‘augmented reality breaching experiment’.

Disappointingly, the breaching experiment conducted at Itchen Abbas did not appear to stimulate a large amount of conversation between the archaeologists. One possible explanation is that because the archaeologists were briefed that a study was being conducted at the field school, they felt asking questions would interrupt or influence the study (the demand effect, see Orne 1962). Another possible explanation is that the students felt they could not engage with the study without appearing as if they were not working, as described previously in the field study. However, it was interesting to note that the vast majority of markers placed through the study indicated the locations where the supervisors and the director had stopped to discuss something with the students while they worked in the trench. The results of the study appear to suggest that the higher up an archaeologist was in the site hierarchy, the more mobile they were in terms of their movements around the site.

3. Implications for Design

The section outlines some of the implications for the design of a technology to support fieldwork archaeologists based on the results of the field studies, interviews and the technology probe.

The field studies at Portus and Itchen Abbas highlighted that supervisors generally spend more time on the edge of an excavation area than within. Furthermore, the observations of archaeologists working at Portus highlights that more often than not, interpretation occurs either standing away from the archaeology or at the edges of a trench. Therefore, designing a technological intervention around the edges of a trench is more likely to capture more information that specifically relates to archaeology than other potential methods, such as creating a wearable device.

The studies conducted at the field school at Itchen Abbas suggest that supervisors and students behave differently around both individual trenches and the entire site. Students tend to remain close to the area they have been allocated to work within, whereas supervisors tend to be more mobile. This suggests that a design requiring a user to move from one area of an excavation to another is more likely to be adopted by archaeologists who are positioned higher in the social hierarchy. Additionally, if we take this further and assume that in some way an archaeologist’s movements around the site approximates that archaeologist’s position within the social hierarchy, we might then consider whether it is possible to disrupt the social hierarchy by disrupting the movements of archaeologists around a site.

The time that archaeologists were observed to spend away from their own trench while talking with colleagues and looking into other trenches at Portus might suggest that having an awareness of the entire archaeology of the site is something that archaeologists feel is both desirable and useful. Designing a technology that achieves this awareness for archaeologists who are generally not as mobile, such as students at a field school or archaeologists working on a commercial excavation, could be an interesting design proposition.

Based on the study at Portus using Looxcie cameras, it is suggested that a technological intervention which captures data about archaeologists, whether visual, audio or gestural data, should be accepted by the archaeologists to the point at which they behave naturally in its
presence. Moving the technology away from a conspicuous location on the body, and perhaps to a device located on the edge of the trench, may help in this regard.

Finally, the field study at Pompeii suggests that deploying any kind of novel technology will likely stimulate new conversation. This has implications for both a technological intervention, and the evaluation of that intervention. An intervention that creates a point of reference for a conversation, such as a visualisation, will clearly have a beneficial impact on the conversation across the excavation. However, when evaluating the intervention, it will be important to minimise or control for how much ‘novelty’ rather than other factors influences the quantity of conversation.

4. Further Research

One way we plan to expand on this research is by conducting further field studies. In particular, we are interested in conducting a field study and repeating the breaching experiment at a commercial excavation, which we feel would provide an interesting counterpoint to the results of the field studies conducted in an academic context and at the field school.

Following the success of this early research, we are also planning an intervention that will explore the themes of communication and disruption further. Our concept is to design and build a device that will capture the ephemeral conversations that occur between archaeologists on the edges of trenches. This device would also have the capability to visualise the content of these conversations to other archaeologists, who may or may not have been included in the original conversation (Fig. 2). It is hoped that such a device would raise an awareness of activity on site, stimulate new communication between archaeologists and disrupt the traditional flow of information supported by the social hierarchy.

References


OpenArchaeoSurvey, or ‘Being Educated by the Digital Fieldwork Assistant’

Jitte Waagen, Nils de Reus and Rogier Kalkers
VU University Amsterdam, The Netherlands

Abstract:
The practice of using mobile survey applications (or a digital fieldwork assistant, dFA) has a tradition of more than a decade in the context of archaeological field survey. In their 2002 CAA paper “Educating the Digital Fieldwork Assistant”, Martijn van Leusen and Nick Ryan wrote extensively about the advantages and practicalities of using a digital field notebook combined with a GPS receiver for field surveys. The OpenArchaeoSurvey project is aimed at improving such applications, building on recent developments in mobile technology. The ‘open’ in our project stands for ‘open source software’, but also for allowing real-time data exchange and communication using the fieldwork application. In addition to the practical advantages, this creates the possibility for all participants to reflect on the collected data. Therefore, the development of the OpenArchaeoSurvey goes hand in hand with a tentative exploration of improved possibilities for Mobile Learning, or, how students can be ‘educated by the fieldwork assistant’.

Keywords:
Field Survey, Mobile Technology, Mobile Learning

1. Introduction

The practice of using a mobile survey application, also known as ‘digital fieldwork assistant’ (dFA) has a tradition of more than two decades in the context of archaeological field survey (introduced in Pascoe et al. 1998, Ryan et al. 1998 and Ryan et al. 1999a/b). In their 2002 CAA paper “Educating the Digital Fieldwork Assistant”, Martijn van Leusen and Nick Ryan wrote extensively about the advantages and practicalities of using a digital field notebook combined with a GPS receiver for archaeological field surveys (Leusen and Ryan 2002). The benefits are related partly to the potential of using GPS locations projected on a map, thus simplifying navigation and facilitating more accurate mapping, and partly to the functionality of a mobile device with GIS technology, being able to bring a large amount of contextual data with you in the field, and allowing for efficient recording of archaeological information (see also Tripcevich 2004). The advantages of mobile technology have also been acknowledged outside the field of archaeology, predominantly in a variety of geospatial and environmental sciences (Wagtendonk and De Jeu 2007, 651). In general, the potential of mobile computing methods is optimal when the following factors are applicable to your research practice (freely adapted from Wagtendonk and De Jeu 2007, 652): the importance of digital data used in the office workflow and the desired speed of data acquisition; the importance of exact field locations; the number of repetitive field measurements; the importance of revisiting measurement locations; the importance of digital field analysis; the importance of objective data collection; direct error control and validation; and the need for real-time information in the field. These factors are all relevant for archaeological field survey, which largely explains the success of digital fieldwork applications in the discipline. The last few years have seen developments, such as the rise in powerful portable computing devices, integrated GPS receivers and affordable 3G network connections, in combination with the maturing of open source software solutions, which offer a huge potential to improve the utilization of the dFA. The OpenArchaeoSurvey project aims to explore this potential. To illustrate where our project takes off, we will start with a short review of the history of the dFA in archaeology. We will then shed light on the improved mobile functionality of our application and explain our experiments in using mobile technology to...
improve an important aspect of archaeological field surveys: educating students. Finally, we will set out our plans for the coming phase of the project.

2. Early Life of the Digital Assistant

In order to understand how the OpenArchaeoSurvey project aims to improve the ‘traditional’ dFA, a short history of the digital assistant in archaeology is presented below.

2.1 A Context-Aware Notebook

The early endeavours in using mobile technology in the field were mainly oriented towards using context-awareness in order to enhance information retrieval and storage (e.g. Ryan et al. 1998). Computer and/or GPS recorded information such as location, time, temperature or user identity could be used to make important notes about a specific location pop-up, for example, and simultaneously to tag collected data. This FieldNote system, intended to work on an Apple Newton device, consisted of a set of modules that under different tasks allowed location tracking, the taking of notes and the display of information (Fig. 1). The application was designed to notify the user, when approaching a location, about which important information was stored in the system. This pre-loaded information could be visualized as either forms or dots on a map. Furthermore, the application allowed for real-time collection of data. The system was designed to run on cheap and lightweight fieldwork computers, to be easy to learn and to avoid distracting the user (Leusen and Ryan 2002, 15). The setup was tested extensively during the University of Groningen Sibaritide 2000 campaign, where its capabilities were successfully utilized in a site-revisiting programme. The main advantages were that the system provided means for an immediate comparison of information collected during previous campaigns, but, most importantly, it was successful in efficiently performing “typical and frequently occurring fieldwork tasks” (Leusen and Ryan 2002, 15). The main vulnerability of the FieldNote system turned out to be GPS accuracy, which is highly dependent on the availability of satellite signals and their quality, and can be a problem in rugged and high relief areas.

2.2 Lightweight Mobile GIS: ArcPad

In the early 2000s, development of an archaeology-specific dFA took a step forward with the adoption of ESRI’s ArcPad as a mobile GIS platform. One example is the archeosurvey application, developed by the Spatial Information Laboratory (SpinLab) of VU University Amsterdam. ArcPad, a lightweight GIS application, offered techniques known from desktop GIS such as layers, map symbology and inquiry tools, but had a compacted user interface optimized for a Personal Digital Assistant (PDA) screen running Windows Mobile (Fig. 2). The ArcPad Application Builder facilitates the production of custom plug-ins to expand functionality, which was used to develop the archeosurvey application.

The archeosurvey plug-in introduced some of the improvements already mentioned above in relation to the FieldNote application (Leusen and Ryan 2002, 13-14), that is, a preconfigured setup (defining pre-loaded data and geographic extent) and on-screen mapping (digitizing polygons). One of the most
significant features of the application was the customizable recording form that popped up after digitizing a feature. This tabbed form featured expanded functionality in terms of form validation, required fields and pull-down lists. Data was stored directly in the attribute table of a shapefile (.shp) and updated in real-time on the screen. This allowed improved control over data quality in terms of completeness, accuracy and consistency. Using this data format rather than the HTML files used by FieldNote permitted its use with a range of desktop GIS programs, increasing the portability of the data. Location information was provided by an external Bluetooth (BT) GPS receiver. BT communication was automatically handled by Windows Mobile, and ArcPad dealt out-of-the-box with the GPS information. The wireless, self-powered BT GPS unit turned the field kit into a more flexible tool for a full day of field walking.

An extensive evaluation of this application showed clear advantages over traditional paper-based methods. The main effects were the effectiveness of navigation, an increase in spatial accuracy, efficiency of data processing in the field and in the office, and the reduction of time spent on error-checking (for the full report see Wagtendonk and De Jeu 2007). Additionally, the use of shapefiles to create an easily updatable GIS database facilitated data flow in the fieldwork project and allowed for teams to take data immediately back into the field for further reference and/or updating. Finally, in relation to the problems concerning GPS accuracy, using multiple high-resolution, compressed raster images as reference material relieved the problem of the GPS being the single source of location information. In general, detailed georeferenced maps and/or aerial or satellite photographs could be used to identify and digitize the contours of a sample unit, with GPS only used for general navigation or where reference material was not sufficient.

The main drawbacks of the archeosurvey application recognized in the evaluation were the requirements in terms of cost and expertise. The time and skill needed to set up and support the PDA with the archeosurvey application remained a clear disadvantage (Wagtendonk and De Jeu 2007, 657).

3. The OpenArchaeoSurvey Project

As mentioned above, the OpenArchaeoSurvey project is aimed at improving the dFA by making use of modern mobile technology. The latter can be characterized by three main developments: operation of smart devices, the adoption of free open source software and the functional benefits of the internet due to affordable 3G data services.

3.1 Smart Devices

The latest generations of mobile devices clearly offer a vast range of new possibilities for mobile fieldwork applications. Netbooks, smartphones and a new generation of tablet-pc’s have dramatically increased the variety in choice of lightweight and high-powered portable computers. Regular fieldwork equipment such as GPS receivers and cameras are regularly built-in to these devices and most offer the necessary technology for a 3G connection. Modern tablets, with their compact form, large screens and integrated functionality,
may be seen as a logical point of departure for an improved dFA. As the OpenArchaeoSurvey project started before their introduction, a different, temporary solution was needed. A netbook with an integrated camera, 3G modem and a collapsible screen, turning it into a tablet PC, a so-called netvertible, was used as a development and testing platform. This device, in combination with a USB GPS receiver, although not particularly optimized for outdoor use, appeared flexible enough to represent the new generation of mobile devices. Moreover, with an unlocked bootloader, it offered the opportunity to install an operating system (OS) of choice.

3.2 Free Open Source Software (FOSS)

A major innovation of the OpenArchaeoSurvey project lies in the choice to move from proprietary software, Windows Mobile and ArcPad, to a free open source environment. This choice was motivated by a series of practical arguments.

In the short term, we realized that using an internet connection would prove problematic with ArcPad, since it does not allow communication with an external port. In the long term, we wanted to minimize costs and be rid of any potential hindrances to the free choice of device and platform, and avoid the so-called ‘vendor lock-in’ (see Weber 2005 for vendor lock-in business models). Furthermore, the general adherence of FOSS to open standards ensures that file formats will be widely usable by a range of software programs and services, increasing the compatibility of the OpenArchaeoSurvey application.

On a more ideological level, though still very practical, was the decision to release the OpenArchaeoSurvey application as FOSS itself. Not only do we have no reason to make it a proprietary and/or closed software application, but open sourcing it may encourage others both to use it and contribute to it, which is likely to increase versatility and longevity.

To maximize our possibilities, Linux, being the OS of choice of a huge range of FOSS packages and surrounded by an enthusiastic user community, was chosen as the development environment for the OpenArchaeoSurvey application. At the start of development, Android, a Linux-based operating system (OS), was rising as a mobile platform at the cost of Windows and iOS. This dominance of Linux-derived operating systems in the mobile market strengthened our expectation that by developing for a very base-standard Linux platform we would not end up too far from code compatibility with at least one of these operating systems by the time their market shares stabilized – whether that was going to be Maemo, MeeGo, Tizen or Android.

As for the GIS environment, we chose Quantum GIS, which is a FOSS alternative to professional proprietary GIS suites. The main benefits of this to our project were the ability to work with a relational database back-end and its extensible plug-in architecture. Comparable to the archesurvey application, QGIS provides common GIS functionality essential for the application, a Python plug-in takes care of customized behaviour. The user-controlled functions of the plug-in can be accessed through a toolbar, presenting a series of buttons to access its features (Fig. 3). A large part of these are basically a rewrite of the classic tools of the dFA, such as capturing location information from the GPS and projecting it as a cursor on a map, on-screen mapping of collection units and data-entry using a customizable data form. Some features, however, have been explicitly developed for the use of an internet connection.

3.3 3G Data Services

It was envisioned that the main advantage of a 3G connection would be the use of a central database stored on a server accessible to all teams and specialists in a fieldwork project (a potential improvement already mentioned by Ryan et al. 1998, Leusen and Ryan 2002 and Wagendonk and Reus 2004). This would
not only allow for easy data management (i.e. no more issues with diverging datasets, automated back-ups), but also enable anyone in the project real-time access to that data (an approach tested using a WIFI setup by Hall and Gray 2004). This would potentially improve the effectiveness of data collection procedures, especially if every specialist could reflect on this data continuously. Furthermore, using a central database on a server would allow access to maps and other reference material that might be useful in the field when encountering unexpected or enigmatic/puzzling archaeological features.

Because a 3G connection without disruptions cannot be guaranteed in remote areas, the OpenArchaeoSurvey application is designed to work with a local dataset consisting of a SQLite database with a SpatiaLite extension. Using the OpenArchaeoSurvey toolbar, the local dataset can be synchronized with the PostgreSQL/PostGIS database on the server at any time a 3G connection is available. This synchronization is record-based and will upload new and/or existing data and download data in a specified time-frame (Fig. 4). Updating the data on the server takes place using a strict policy that avoids overwriting newer records with older information. In the case of downloading data, the application simply downloads and replaces all records in the local database and immediately refreshes the map view.

To enhance the potential of sharing data during fieldwork, two more features were developed. First, the internal camera of the mobile device was integrated into the application. Pictures can be taken using a button on the OpenArchaeoSurvey toolbar that accesses the camera, and the GPS information can be used to anchor it to the location where it was taken. The photograph then appears as a location on the map, showing a thumbnail when clicked. These photographs can be exchanged with other teams and specialists using the central server. Second, a chat-client was integrated into the QGIS layout which connects automatically to a preconfigured IRC server, enabling users of the application to discuss findings and provide direct feedback.

A specific advantage of 3G network access is, evidently, the ability to consult online resources. Although implementation of a website as part of the OpenArchaeoSurvey project was initially modest, a community site has been set up. The use of the site has been twofold. Initially it acted as a back-up for use in the field should the OpenArchaeoSurvey application breakdown, with the community site featuring a chat room and the possibility to up- and download files. Subsequently, a logical next step seemed to be to develop this website into an online community using a forum and a wiki (with a manual and a trouble-shooter for the software), providing online resources for users in the field.
4. Being Educated by the Fieldwork Assistant

As may be clear from the application description, the ‘open’ in our project not only stands for FOSS but also reflects the idea of opening up the digital data to all participants in the project using the fieldwork application. Apart from the practical benefits, sharing data, combined with the possibility of experts reflecting on that data, opens up possibilities for education. Therefore, the development of the OpenArchaeoSurvey project goes hand in hand with a tentative exploration of improved possibilities for Mobile Learning, or, how students can be ‘educated by the fieldwork assistant’.

Our project included a few tests, executed as fieldwork pilots, to see what the additional benefit might be with respect to educating students during fieldwork. While the OpenArchaeoSurvey application was developed within a broader project aimed at improving education through IT, we were no experts in ‘mobile learning’. Nonetheless, we attempted to intuitively and experimentally apply basic concepts of mobile learning in our project setting.

4.1 Mobile Learning Principles in a Nutshell

A point of departure was the Manolo Project, which was aimed at exploring the possibilities of mobile learning and led to the development of the initial archeosurvey application (see Wentzel et al. 2005 for an introduction). Mainly oriented according to the ‘anytime-anywhere’ paradigm in mobile learning (unlimited access to learning materials), the project hoped to increase the learning impact by bringing the classroom to field locations where the objects of study were at hand. This ‘virtual classroom’ consisted of learning materials provided on mobile devices and/or communication with peers and teachers (see also Armstrong and Bennet 2005). As the learning materials were rarely actively assessed by students in the field, and a 3G connection was not implemented in the fieldwork setup around the archeosurvey application, actual mobile learning remained relatively modest. The main learning results pertained to students becoming familiar with the concepts of mobile GIS and navigation using GPS, and gaining an understanding of the potential of these techniques in the context of archaeological field survey (Wagtendonk and De Jeu 2007, 659). An indirect, though intentional effect on learning conditions was the reduction in data processing time, allowing the students more time to participate in the scientific evaluation of that data (Wagtendonk and De Jeu 2007, 259). In the end, the archeosurvey application was a modest but firm step forward in increasing the opportunity for education during field surveys.

More recent approaches to mobile learning focus on adapting learning material to user pre-knowledge and learning preferences, called ‘adaptive m-learning’ (Burghardt et al. 2007), which aims to find a compromise between intrinsic learning activities and support from outside. It focuses on adjusting to spatial as well as temporal contexts, considering the learning process of the user under changing conditions. On the practical side, this amounts to capturing user knowledge and learning progress, modelling user activities and context, adjustment of the learning content presented and the evaluation of the mobile learning environment. In the end, the ‘anytime-anywhere’ approach to mobile learning is redefined as “any sort of learning that happens when the learner is not at a fixed, predetermined location, taking advantage of the learning opportunities offered by mobile technologies” (Burghardt et al. 2007). This type of mobile learning is therefore highly dependent on a 3G connection, allowing teachers to monitor and adjust the learning content of assignments.

4.2 The Learning Scenarios

Our attempt at mobile learning could be characterized as an adaptive m-learning
approach, heavily based on continuous monitoring, providing feedback and adjusting assignments for students depending on the data and its context. The difference, however, was that we worked with predefined learning scenarios, a prepared assignment and a specific learning objective, organized in addition to regular surveying. This approach was chosen because the OpenArchaeoSurvey project was an extension of classroom education, which also has distinctive learning objectives.

The first learning scenario was aimed at making students aware of the rationale for recording sample context data. This scenario has been tested in Molise, Italy, during the Sacred Landscape Project and on Zakynthos, Greece, during the Zakynthos Archaeology Project (for recent reports, see Pelgrom and Stek 2010 on the Sacred Landscape Project, Van Wijngaarden et al. 2008 on the Zakynthos Archaeology Project). Students were sent to fields that had been sampled five years earlier, with the assignment to repeat the data collection. The result was exchanged with the GIS expert in the field office, who produced weighted density maps and compared results with the maps of five years ago. The outcome was sent back to the students, who were challenged to explain differences by looking at find circumstances, recent agricultural activity, etc. While this was potentially a nice scenario and executed to some success, the learning effect was compromised by the complexity of setting up the scenario. It proved quite difficult to find a real-world situation (e.g. a readily accessible field, preferably with changed agricultural conditions) in which the data and context observations actually made useful sense (e.g. the results would not differ much).

The second learning scenario revolved around site revisits. This scenario has been tested in Thessaly, Greece, during the Halos Archaeology Project (for an introduction to the project see Reinders 1998). Students were equipped with georeferenced maps showing site locations surveyed a decade ago. They were asked to retrace a site, draw its boundaries and collect data using clearly defined criteria. After uploading their results they would then receive immediate feedback from pottery specialists and a GIS/methodology expert on their decisions and possibly be asked to redefine criteria and data collection strategies. Again, although the scenario was executed to some success, in many cases the difficulties of retracing the sites alone turned the assignment into a complex undertaking. The intervention of experts was often needed much earlier, before the actual assignment could be executed.

Though the use of the learning scenarios was not without merits, they often turned out to be too complex to put into practice. However, it became clear that merely staging and continually monitoring simple revisits with the OpenArchaeoSurvey application provoked the participants into asking questions. The exchange of information amounted to a discussion between students (e.g. about the functionality of the application or GPS navigation), as well as requests for feedback on information or expert opinions (e.g. What does the paper report say on site x? or What dimensions should feature y typically have?). It appeared that by enabling continuous interaction and data exchange, the application could function as a learning-community builder. The learning was directed by the specific circumstances and problems the teams confronted during their task, and they benefited by gaining insight into the decisions and opinions of experts or senior staff members. In our view, this appears to reflect Burghardt’s conclusion that “m-learning has great potential to support the independent formation of knowledge by exploration” (Burghardt et al. 2007). Other than learning effects, the continuous exchange of information and opinions showed the potential to increase the efficiency of fieldwork, simply because the teams were ‘smarter’ and more involved. Based on these conclusions, our aim for future deployment of the application will be to refine an approach of defining education objectives that expand upon this type of mobile learning.
5. Future Prospects

Apart from aspects of mobile learning, the OpenArchaeoSurvey application is currently still under development and we have some specific improvements planned. Furthermore, the evaluations have led us to focus on a series of features that are likely to be our future fields of study and development.

5.1 Development Platform

Recently it has become clear that the Linux kernel will be updated with the Android code produced by Google, effectively merging the Linux and Android kernels, considerably increasing code compatibility for Linux and Android applications (Corbet 2011). For the OpenArchaeoSurvey project specifically, this means that the prospect of Quantum GIS running on Android with a functional Python run-time environment will become probable, which would make it run on out-of-the-box Android devices somewhere in the foreseeable future.

As for the code in its current state, we are finalizing the documentation of OpenArchaeoSurvey version 1.0. The code will be available via the code hosting site Bitbucket. The release of the application will be announced on the community website and accompanied by an open invitation for everyone interested to download, test and use the application, as well as participate in its development.

As mentioned above, tablets are becoming increasingly of interest to us, and currently we are staging a pilot using the application on a tablet PC for the 2012 summer fieldwork season. Moreover, we eagerly await the arrival of the first consumer tablets with PixelQi screens, due to their superior outdoor readable screen technology.

5.2 Functionality

While currently not part of planned improvements to the OpenArchaeoSurvey, below we list what we consider to be beneficial enhancements regarding functionality.

- Integration of a finds-processing database

Having access in the field to data available in a finds-processing database could be very useful, all the more when combined with visual representation of that material. This would offer the potential to assess data that may answer questions that come up during fieldwork and will make the processing of new finds even more efficient. Finally, having examples at hand for reference, it is likely to improve the accuracy of determining newly collected material.

- WMS

The option to access maps using a web map service (WMS) would be relatively easy to implement, which is a feature already available in QGIS. Maps could be retrieved from existing databases as well as a database built with project-specific material using an application such as MapServer. Of course, this option would only be beneficial if the mobile device itself was unfit to store a spatial database of considerable size. An interesting use would be to have access to old aerial photographs when encountering unexpected or puzzling/enigmatic archaeological features.

- Integration of other data recording features/equipment

In addition to pictures, voice notes could be relatively easily recorded and georeferenced (experimented with by Tripcevich 2004, 21). Using a wireless network or BT connectivity, output from other devices could be integrated into the spatial database of the OpenArchaeoSurvey application. Some options have already been noted, for example, barcode readers for finds processing (Leusen and Ryan 2002, 12), though additional cameras, remote-sensing equipment or even the traditional clicker would most likely also prove to be useful.
improvements (Tripcevich also notes digital callipers and scales, Tripcevich 2004, 17). Not only would this increase efficiency and accuracy, it would also make teams more flexible as data could be collected by several team members.

- Augmented reality

When access to the internet through 3G or 4G networks becomes truly reliable and fast enough in remote areas, real-time projection of data available onto the internet will become a realistic option. Maps and forms could be augmented with data from, for example, Wikipedia, Flickr and/or archaeology-specific services.

5.3 Support

An intended effect of the project is that the application may be used by other parties. In addition to archaeologists, biologists, geoscientists and social scientists can also apply fieldwork settings for learning and data gathering processes. As previously mentioned, one of the main problems with these kinds of applications is the high level of investment in expertise and equipment (Wagtendonk and De Jeu 2007, 657). In the current phase of the project, we are concentrating on this problem. The main focus will be the website (see below) where we will offer materials to support the setup and use of the OpenArchaeoSurvey application. We plan to record screencasts to assist others through the necessary steps, we will host a manual and necessary documentation and open a forum for discussion and FAQ. On the technical side, we hope to be able to produce a virtual appliance that will help users to set up a preconfigured server. We will also make the forms used in the application compliant with the Quasar Toolkit (Qt), so an existing graphical application can be used to adjust forms for project-specific databases.

5.4 Website

As described above, the website for the OpenArchaeoSurvey project initially concerned features supporting the application and its infield use. However, given the emphasis of the project on improving learning possibilities, the site was quickly deployed for secondary support. Course syllabi were processed into wiki pages to provide background material on survey methodology and computational archaeology to fieldwork students. To allow for interaction and building an FAQ, a forum was set up for discussion of more general topics. With the realization that such a knowledge-base and communication platform would have merits beyond the OpenArchaeoSurvey project, it became an entirely new e-learning project in itself (see Waagen et al. 2012).

6. Conclusions

Technological developments over the past decade have opened up a wide range of potential improvements for the traditional digital fieldwork assistant used in archaeological field survey. We focused on three areas of modern technology to explore this potential: modern mobile devices, free open source software and 3G networks. Additionally, enabling data exchange and direct communication, we tentatively experimented with the potential for applying mobile learning during fieldwork. To conclude, we would like to sum up some of the preliminary results of the OpenArchaeoSurvey project. They are necessarily of a qualitative nature because, due to the emphasis of the project on mobile learning, its practical and organizational benefits over the archeosurvey application have not been evaluated in measurable units. Since the pilots have been of limited scale (one team) and the learning effects cannot be simply measured in the short term, quantitative data is not yet available. However, adding to the findings already mentioned above, below we present a summary of the main advantages and disadvantages.

6.1 Disadvantages

Most disadvantages relate to practical issues which primarily concern the use of a
device that is not yet optimized for rugged outdoor use. Furthermore, setting up the hardware and software requires a relatively high level of technical knowledge and can be rather time-consuming. Finally, learning scenarios developed for survey fieldwork tend to be difficult to put into practice.

6.2 Advantages

In addition to the benefits of the pre-existing archeosurvey ArcPad application mentioned in the paper, OpenArcheoSurvey application allows users to choose a lightweight, powerful integrated device (from a wide range of types and brands). The software is free open source, increasing its potential longevity and versatility and making it available to any interested party. It can increase the efficiency of both data management and a project’s workflow in general and it demonstrates interesting potential for adaptive mobile learning.

Finally, ten years ago we were ‘educating the digital fieldwork assistant’, now we are developing software that provides the opportunity to ‘be educated by it’. Apart from the practical benefits, we believe that archaeological field surveys can gain much from mobile learning: if all participants are involved and better informed this encourages more enthusiastic participation, in turn improving data quality and the educational potential of a fieldwork project.

Acknowledgements

We would like to express our gratitude to Dr. Martijn van Leusen and Drs. Alfred Wagendorf for their critical comments on early drafts of this paper. Prof. Dr. Vladimir Stissi and Dr. Gert-Jan van Wijngaarden have been generous to allow us valuable time to experiment with the OpenArcheoSurvey application and their fieldwork crew. We are greatly in debt to the ICT for Education (ICTO) programme at the University of Amsterdam for making this project possible. The application was developed in collaboration with the SpinLab of VU University Amsterdam. Special thanks go to Joeri Wetters for the graphic design of the toolbar.

Disclaimer

The OpenArcheoSurvey project formerly bore the name Learning Sites, and was presented as such at the 2012 CAA in Southampton. However, to avoid confusion with an already existing American virtual heritage company called Learning Sites we changed our name. Hereby we state explicitly that there is not, nor has there been, any relationship between our OpenArcheoSurvey project, formerly Learning Sites, and the Learning Sites company, Williamstown, USA.

References


personal technology for the field.” *Personal Technologies* 2: 28-36.


The Use of iPad as a Documenting Tool on an Archaeological Excavation on Govče 2011 Project in North - Eastern Slovenia

Eva Butina
Institute for the Protection of Cultural Heritage of Slovenia, Slovenia

Abstract:
The paper presents the use of a tablet PC on an archaeological excavation in Slovenia conducted by the Centre for preventive archaeology (CPA). Standard Apple iPad2 was used to support archaeological documentation on site and to provide quick and easy mean of managing of gathered field data with the possibility to manipulate and create archaeological data during the excavation on site. The use of the tablet on the site provided a very useful tool for archaeologist for fast and accurate documenting of specific features like skeleton remains and specific small find in graves excavated. Since the site was and is still being excavated in a remote area in hard environmental conditions (early winter and hot summer months), the use of a tablet PC on-site has proved to be a practical tool and a big advantage for archaeologists.

Keywords:
Tablet Computer, Documentation, Excavation Site

1. Introduction

The majority of present day archaeological excavation in Slovenia is the so called rescue archaeological excavation in protected archaeological sites which are planned for development. The rescue archaeological excavations represent about a third of all archaeological field research in Slovenia. With the new law which was passed in 2008 concerning the cultural (archaeological) heritage of Slovenia, a Centre for preventive archaeology (CPA) was created in 2009 as part of The Institute for the protection of cultural heritage of Slovenia. In its purpose, CPA is very similar to French INRAP: to research a protected or a potential archaeological site before the actual development takes place.

At the Govče 2011 project the CPA archaeological team excavated a part of a cultural heritage site - a late 16th century protestant church and the adjoined graveyard. The use of a tablet computer - iPad was introduced as a possible technical support for field documenting since there was no printer on site in winter season, and later in the summer season as an auxiliary form of documenting the site. The use of the tablet computer in the Govče 2011 project was not planed in advance and this paper presents experience with the use of a tablet computer on an excavation site without prior experience with tablet devices and in this case iOS.

The first part of the paper presents archaeological system of documentation used by CPA which is necessary in order to present limitations in the use of a tablet computer as a documentation tool for this archaeological excavation and other field research conducted by CPA.

The second part of the paper presents the use of the iPad on the Govče 2011 project and future possibilities of tablet devices on archaeological field research projects.

2. Archaeological System of Documentation

Combining impossibly short time tables with low budget and very often extreme weather conditions is a big challenge for every
The Use of iPad as a Documenting Tool on an Archaeological Excavation on Govče 2011 Project

Eva Butina

field archaeologist. Dealing with the enormous amount of data generated by a stratigraphic method of single context recording, the use of a computer on site has become a must in order to produce an accurate and coherent archaeological documentation of the site as demanded by the archaeological standards. In January 2011 CPA issued internal Minimal standards of archaeological documentation of all archaeological research which include specific description of methodology for field and underwater excavation, field survey and cabinet analysis of larger areas, gathering and managing of archaeological filed data and the organization of field (digital and paper) documentation, the so called digital archive of an project.

The minimal standards for archaeological excavation used by CPA are:

1. The use of total station for spatial documenting of:
   - all stratigraphic units on site (single context documenting) and the creation of a site plan in a CAD drawing (DWG file) with the program module Skavt®,
   - illustrative cross-sections of the site and of specific contexts,
   - “photosketches” of specific contexts and situations; a photosketch is a rectified digital orthophoto which serves as a practical graphical lining for an archaeologist to interpret specific details of a photographed archaeological feature (Fig. 1). The interpretation (definition of materials, SU numbers and their stratigraphic relations, small find and samples ID numbers, topographic details...) is made on-site, using water resistant markers on a PVC foil which is laid over the printed photograph. These on-site interpreted photosketches are necessary for later accurate detail vectorisation in a CAD drawing (Fig. 2), since many of the details can’t be observed or distinguished even on a high resolution digital photograph. The technic of photosketches is a faster and more accurate way of producing archaeological drawings of plans or cross-sections of the site,
   - various samples, small finds, skeletons etc.

2. The use of digital camera for all photographic documenting on the site.

3. The use of prescribed CPA paper forms for field documentation (descriptions of stratigraphic units and skeletons, lists of finds, samples, drawings...).

4. The organization of the project’s digital archive (structure and naming of folders

---

2 Program module Skavt® was specially developed for processing and managing of archaeological data gathered on an archaeological excavation. The program module works on the AutoCAD platform. It enables processing of gathered data (transformation of spatial data gathered with total station into CAD objects), fast and accurate creation of site plan and digital drawings of specific features. It also provides archaeologists with an easy to use tool for searching specific entities within the site plan and to input observational data or creating additional data for each entity.
and files) is predetermined as it is formed automatically by the program module Skavt® when the project is started.

The defined organization of the digital archive of all archaeological projects is important for later archiving and easy searching for specific data in the archive by people who were not involved with the project directly.

The most important part of the archive is the folder Plans where the site plan is located together with all the data connecting to the spatial data gathered with the total station: the measurements in the form of IDX or SDR file, CAD drawing of every stratigraphic unit, feature or cross-section, rectified digital ortophotos and cartographic images (DOF, TTN) imported into CAD drawings as external references, vectorisations of Photosketches, other data that can be imported or exported using Skavt® into/from the site plan.

This flexible system of archaeological documentation has been developed thru practical employment in the field for more than a decade and allows for an accurate and fast gathering of archaeological field data even with a bare minimum depending on the size of the site. For a site of circa 200 m² a research team of at least two people (not including the hired manual labor force) is needed equipped with a total station, a digital photo camera, a laptop computer with CAD software and module Skavt®. The system of documentation is so well established and practically tested in the field that allows for relatively easy workflow even for a team with almost no prior experience working with a total station or CAD computer software since Skavt® also includes a detailed tutorial for every operation in processing and managing of gathered data.

The system as described above is electrically independent since all of the mentioned instruments operate on batteries and therefore is suitable for remote areas with no electrical supply. But to work comfortably, archaeologists need a color printer on site in order to provide the team with up to date information of the situation at the site. This is done by printing 2D plans of daily situations and of the photosketches for on-site interpretation of the details. On the numerous field excavations during the years it became clear that these daily printouts of last spatially recorded situations are vital for the team to check for any mistakes or changes made during spatial documenting of context, small finds etc. Daily printouts are necessary since constantly checking the situation on the computer during excavation is not practical (the computer preferably being located in a location secure from the elements). It is more efficient to make comments and sketches directly on the printouts on-site and then to provide these updated situations with comments to the person responsible for managing the site plan as soon as possible, so that he or she can make the proper adjustments in the site plan.

2. The Use of iPad on a Govče 2011 Project

We had an opportunity of using an iPad on an archaeological excavation in Govče, a remote
area around 60 km north-west from the capital of Ljubljana. The excavation site of around 2100 m² was situated right next to a protected archaeological site of a former protestant church from the end of the 16th century. The excavations have begun in November 2011, when a portion of the site (roughly a third) has been excavated, revealing a large layer of stone, brick and mortar, and some modern day post holes for agricultural purposes. In mid May 2012, the excavations resumed and were finished on time by the end of July 2012. Remains of church circular wall foundations were unearthed which is consistent with the historic encounters of the site, together with 31 burials, with skeleton remains in very poor condition since many of the grave pits were dug under the level of groundwater.

When the excavation first begun in November 2011, there were specific circumstances that enforced the use of an alternative documenting tool on site. A small team of two archaeologists and one archaeological technician were facing a plan to excavate an area of circa 400 m² in only 10 days before the winter comes. The weather conditions were very bad, with temperature dropping below -6°C with an average of around 0°C. The winter sun also created long shadows that provided very difficult conditions for photographic documenting of the site and contributed to even more difficult documenting of the frozen stratigraphic layers.

Technical equipment included a total station, a digital camera and a laptop computer. There was no printer on site which made our system of documenting incomplete without the ability to provide the photosketches for on-site interpretation. These on-site interpreted photosketches are standard norm of documenting and with not enough time or manpower for making manual scale drawings, we had to come up with a solution in order to finish the task of excavating the desired area in time. Since constantly going to the nearest town for printing was not an option, we came up with an idea of using of one of the team member's tablet computer.

The device, an iPad2 with 16 GB of storage, was fairly new and none of the team members had much experience with tablet computers. But since the device enabled a quick caption of photographs and easy adding of comment on the captured pictures by scribbling or drawing directly onto them (interpretation of the photographed scene) on-site, it soon became apparent that this is a very good and possibly more practical substitute for our existing technique of photosketches. There are many applications for sketching and drawing available today. We would have probably been able to produce digital technical drawings fairly easily using applications such as iDraw like they did on Pompei project (Wallrodt 2011) or E’se’get Archaeology Project (Betts 2012), but these would require more of valuable field time and the use of a stylus which we didn’t have. Archaeological mapping or plan drawing using a tablet computer takes the same amount of time as standard drawing with paper and pencil (Betts 2012), compared to the photosketches technique which requires much less time on-site and therefore the field work can proceed more quickly. For these reasons and because we had limited time and resources in the field but not in the office, we used the application called SketchBook for making quick and easy on-site sketches of plans and cross-sections and interpretations of some of the small sized photosketches. Later in the office we had enough time to use these digitally interpreted photosketches combined with spatial measurements and rectified orthophotos for producing scaled vector drawings of plans and cross-sections of the site.

During these first ten days of the excavation, we were able to try out some other features of the device with the help of numerous practical software applications for tablets like viewing the CAD drawings of plans and sections of the site, viewing high resolution photographs in the photo archive, inputting some of the data
directly into our digital documentation with existing folders and documents and even send data for viewing and backup via wi–fi (in the nearest town) or via the smart phone connection right on-site. All this was done without any prior change to the existing documents in the digital archive of the project.

The Autodesk software AutoCAD WS is a free software application for tablets and android devices that allowed us not only viewing of the DWG file of the site plan, but also enabled to edit the plan with the basic functions of the AutoCAD software (adding, moving, deleting, copying, scaling objects or text in the CAD drawing). This allows for on-site management of site plan while discussing specific situations with team members or for simple and fast acquirement of special data considering specific SU or feature that has already been documented. But since the AutoCAD WS doesn’t support programming, we were not able to use our program module Skavt® for managing hundreds of layers in the site plan which made it unpractical for any advanced use or management of the plan. The other deficiency was the inability of viewing the rectified and georeferenced photo sketches in the site plan.

With resuming of the excavation at Govče in May 2012, the team was enhanced by additional team members: a total station operator, a photographer and two extra archaeological technicians. There was also an additional laptop computer for sorting and editing the growing photo archive, and a printer. We decided to again use the iPad as a documenting tool on site as an additional technical tool since we were not able to bring the device on the excavation site every day, mainly for the purpose of documenting of everyday excavation process like a digital diary and of course for viewing and managing of the site plan and its digital documentation on-site (digital forms) whenever necessary.

We used several apps for iPad that enabled us to work with existing files (mostly MS Office: Word, Excel and Access files) in the digital archive of the project. For direct input of observational data for purposes of site documentation, we used the Apple application called Pages which enabled us to write observational data directly into the digital form of prescribed paper forms because we didn’t have a functional database for digitally recorded observational data at the time. Besides we only had one device on the excavation site and therefore we only recorded some of the stratigraphic units digitally, most of the observational data was still made the old school – using paper forms. For making quick and easy on-site sketches of plans and cross-sections and some of the photo sketches (Fig. 3), we used the Autodesk application called SketchBook.

The major breakthrough in the use of the iPad on the excavation site in summer 2012 was the newest release of the AutoCAD WS software
The Use of iPad as a Documenting Tool on an Archaeological Excavation on Govče 2011 Project

Eva Butina

for mobile devices for viewing and management of the site plan. Now the user is able to not only to view, edit and download all the changes made to the DWG drawing, one is able to edit files even without internet connection since the changes will be synchronized as soon as the user gets online. This is a very good way of archiving the site plan (necessary back up). Also the drawing can be easily shared among members of the team on-site (or off-site), and edits and markups are immediately visible to all of them. We can easily use this feature not only to communicate between the team members on-site but also to inform and update the investor on the progress of the excavation project.

The new version of AutoCAD WS also enables simple management of external references (xRef files) within the drawing which was very important for viewing the photosketches and other georeferenced images in the site plan on iPad. But there are some rules to be observed in order to achieve this quickly and properly. The most important thing is to make sure that the actual xRef files have the same relative path as in the original file on our PC or laptop computer otherwise we will have to attach each file to the drawing again, which is almost impossible since AutoCAD WS doesn’t have software modules for correlating imported images.

The other set back in the AutoCAD WS is the draw order of objects in the drawing. It is important to set the draw order of objects before it is uploaded to a tablet or mobile device since this feature is not supported AutoCAD WS. Once the drawing is uploaded, the objects, like for example images, can lay over other objects and entities in the drawing, and these can only be seen, if the layer with the image is turned off (Fig. 4). This makes any use of georeferenced images like DOF, TTN or photo sketches on a tablet useless unless properly ordered in the drawing on the computer beforehand.

3. Potential Use of Tablets on Excavation Sites in the Future

The use of the iPad on the excavation site in Govče 2011 project was borne out of practical need for substituting the lack of printer on the excavation site. The outcome was a pleasant surprise for the team since it turned out that the device is sturdy enough for documenting on site in complex or even hard environmental conditions (minus 6°C and high humidity in winter, and plus 36°C with fine dust everywhere in early summer). The iPad we used on the excavation site didn’t have any additional protective cover apart from the basic smart cover for the touchscreen. The device operated with no problem for over 8 hours every day compared to the laptops which regularly overheated during the hot summer days and were useless after two hours of working without the electricity supply from the generator on the site.

The device we used worked faster than the laptop (i5 CPU 2,4 GHz, 4GB ram, x64 OS) that was used as the main processing tool for the gathered data, especially while browsing through the photo archive or viewing objects in 2D and especially 3D in DWG drawing, and was very easy to handle (Forstnerič 2011).

We were able to view and edit the existing data in the digital archive of the project once it was uploaded to the device. The actual
The uploading of the files and folders was the more difficult part for the team (MS OS users), since one can only do that by using Apple software iTunes which took some getting used to. Once we had the mobile device synchronized with a computer, the actual editing of the files on the device was pretty much straightforward.

The other very practical feature is the simple downloading and uploading of desired files like DWG drawing, photographs or reports onto the cloud directly from the site, using free software for remote storage of data (Google Drive, Microsoft SkyDrive, Dropbox...). This provides a uniquely remarkable way of archiving field data and at the same time sharing the information with other researchers or members of the team as was implemented by the team of the E’se’get Archaeology Project (Betts 2012). The obvious advantage of the use of mobile devices like tablet computers in the field is the easy and fast connectivity to other devices and to the internet. The use of a wi-fi network is the easiest way to share or move data between devices, but this is not always possible, either because there is no wi-fi available on-site or because there are technical difficulties preventing the use of wi-fi, for example under or near the power lines. With the remarkable decline in prices for data transfer via 3G mobile network for smart phones in the past few months, sharing and exchanging data became an affordable practice for field archaeologists in remote areas. All we need is a smart phone with 3G network, a decent GSM signal and a tablet computer. It takes a few minutes and a couple of MB of data transfer every day to provide or find information we need, to archive specific data or simply to correspond via e-mail. On the Govče 2011 project there was no wi-fi network available. When in a place with wi-fi availability, Dropbox was used mainly for the purpose of archiving the excavation site plan (DWG file) and spatial data (IDX or SDR files).

The only obvious deficiency of the tablets currently available on the market (including the iPad used) is the screen. The high resolution touch screen (iPad2 has 8 LED-backlit IPS TFT, capacitive touchscreen, 16M colours, 768 x 1024 pixels, 9.7 inches) is covered with scratch-resistant glass and oleophobic coating that should give a better resistance to oily fingers (GSMArena 2012), has a high gloss glass cover that is not very practical for use in the open space at strong sunlight as it reflects everything and makes very difficult to see what is on screen (Fig. 5). This is especially noticeable when taking pictures with the device: one can only focus on the objects while observing the screen. But we learned on the Govče excavation site, that it just takes some practice and some extra pictures (they can be easily deleted later), just to be sure we did capture the desired scene or object. Although the operating of the device was impaired in strong sunlight, with minimal shading similar to the one when using a smart phone outdoors, we were able to make notes and drawing on pictures taken, input observational data directly into the digital documentation forms and view and edit site plan DWG drawings.

4. Conclusions

There is certainly a bright future for mobile devices on archaeological field research projects. With new application software and on the cloud technology that enables the user to perform (some of) the functions from their
laptops or PCs on their mobile devices, they become more useful every day. Not just for browsing through the net or pictures stored in the device or online, but for serious work also. The new generation mobile devices enable archaeologist to manage the gathered archaeological data in real time, communicate with other team members effortlessly and perform his or hers tasks on site faster and more efficiently. However the processing of large amount of data and creating or editing exact drawing will remain the domain of standard laptop and PC as it requires the use of a mouse for comfortable and precise work. Unless tablets are developed with OS similar to that of PC, that will enable the archaeologist to use the programs and data file formats they are used to, without converting from one format to another when using data on several devices with different OS. The latest news in the digital word sound promising with the imminent release of MS Surface Pro tablet computer (Microsoft 2012).

A tablet computer on an excavation site is a big help for archaeologists. From our experience on the Govče 2011 project, the device proved most useful for on archaeologist overseeing a sector or en entire excavation site, since it enabled to overview all the aspects of the project: measurements with the total station, observational data, pictures of the documented archaeological features and small finds etc. The improved communication between team members by presenting specific data and instructions, together with graphic documentation data in the form of sketches or photographs while standing in the excavation trench, proved to be extremely practical, especially on a large excavation site such as this.

Provided there are several devices on a specific site, there is the possibility of documenting every aspect of archaeological documentation digitally. But that requires a well organised data base and practical digital forms for observational data. If creating a simple digital form and a database doesn’t require a lot of time and knowledge - many archaeologists have successfully created simple systems for inputting and archiving data (Betts 2012), there is the issue of the technical problem of inputting data into site’s database by several users simultaneously. Some of the problems concerning the so called syncing of database from several devices over a network are described in Paperless archaeology blog (Wallrodt 2011).

We used the iPad as an auxiliary tool on site and not as the primary documentation tool because of the issue of all digital documentation and because we were not able to create a database in such a short period of time, that would allow for an easy and fast input and especially output of the desired observational and spatial data gathered on the excavation site in a format as required by the CPA Minimal standards of archaeological documentation. These first practical tests in the field have proved successful and provided with a plan when organizing an excavation project, which includes a tablet computer as an actual tool for field documentation, documenting an archaeological site will be faster and more practical.

References


Back into Pleistocene Waters. 
The Narrative Museum of Casal de’ Pazzi (Rome)

Augusto Palombini  
CNR - Istituto per le Tecnologie Applicate ai Beni Culturali, Italy

Patrizia Gioia  
Comune di Roma, Italy

Antonia Arnoldus-Huyzendveld  
Digiter s.r.l., Italy

Marco Di Ioia and Sofia Pescarin  
CNR- Istituto per le Tecnologie Applicate ai Beni Culturali, Italy

Abstract:  
The Casal de’ Pazzi museum exists since the early 90s on an ancient fluvial deposit of the Aniene river ancient riverbed (at some 200,000 years ago). During the Pleistocene, the site was very different from the actual situation, both from the geo-morphological point of view and from the vegetation and zoological one, as remains of Elephas Antiquus, aurochs, hippos, bears etc. were found, as well as a Neanderthal skull fragment. Since its birth, the Museum followed a strong narrative strategy, so that it was quite easy to plan a digitally-driven improvement. The virtual museum, planned and created in about 2 years, consists in a touch screen educational game for pupils (the Plei-sto-station), and in a movie on Elephas Antiquus and Neanderthal man’s life. Moreover, a projection targeted to the deposit will perform the effect of river refilling.

Keywords:  
Pleistocene, Prehistory, Virtual Museums, Palaeo-environment, Digital Reconstructions

Which is the real core of a virtual museum? What makes the difference when dealing with traditional Cultural Heritage dissemination and digital technologies? A general overview of the many existing digital applications for museums leads to identify the technology as one of the tools to be used in the frame of a wider strategy whose core is the narrative approach (Antinucci 2007, Forte et al. 2006), where such an expression may be someway defined as the complex of story-telling techniques to be used in order to perform a communication as effective as possible of the Cultural Heritage.

The reasons for such a relevance of narration in terms of enhancement of emotional and learning power, is probably related to deep cognitive dynamics in managing the connections among symbols and objects (Forte 2007; Gibson 1979), and this is not the place to analyse it. Nevertheless, following such an idea of story-telling as a key factor, two consequences should be evident:

- technology, as happened for the movie industry, is just a tool to better perform narrative strategies.
- the narrative approach may exist even without technology, as for some museums born before the spread of digital technology or with a low - well targeted - use of technology (Palombini 2012).

The Casal de’ Pazzi Museum is a good example in this direction as it represents, since its birth, an endless effort towards a narrative approach to dissemination, so that the recent
1. The Site

The history of ancient Rome, as documented by the monumental remains of the classical period, is the main interest for many people visiting the city. Nevertheless, many prehistoric and proto-historic sites also exist in the Roman territory. Ten Middle Pleistocene sites are located both in the surroundings of the Italian capital and within the town itself (Anzidei et al. 2001; Anzidei, Gioia, Mussi 2004). The earliest ones are dated to 300 ka BP, three others date back to 200 ka BP, an only one (Saccopastore) is dated to c.100 ka BP. They are all chronologically placed in warm interglacial stages, when the sea-level was higher than now, and the rivers were depositing sediments.

Many of them were found in the lower Aniene valley (Fig. 1). This tributary river flows into the Tiber within modern Rome. Before the Second World War, the territory crossed by the Aniene river was open countryside and numerous Pleistocene deposits, including archaeological sites, still existed at the time: at Saccopastore, two Neanderthal skulls were discovered in the 20s and 30s of the last century (Manzi et al. 2001; Manzi 2004). Due to the city expansion, this area is now fully urbanized. Among these sites, Casal de’ Pazzi is the only surviving site, now well within the new Rome quarters. It was discovered in 1981 (Anzidei 1983), during excavations to build city infrastructures. A 200,000 years old fossil branch of an ancient fluvial deposit was at that time unearthed. After the discovery of the first fossil tusk, works gave way to archaeological excavations, which lasted until 1986 (Fig. 2).

The site is located at an altitude between 30 and 34 m a.s.l. along the lower Aniene valley. There are several geological maps referring to the area, the most recent and most detailed being the 1:10,000 scale map by Funiciello & Giordano (2005). We have locally a rich representation of the Middle to Upper Pleistocene geological events, characterized by the intertwining of volcanic activity, tectonic activity and sea level changes. The formations are represented mainly by Middle Pleistocene deposits, topped stratigraphically by the Upper Pleistocene Saccopastore unit, a fluvial deposit in which the site is incorporated.

In the area in which is located the archaeological site, the Tufo lionato is the volcanic formation outcropping throughout the territory. It dates back to about 360Ka (Karner et al. 2001), and constitutes the bedrock in
which the Aniene river excavated his course during the Pleistocene. The Aniene incised the bed rock, and then started filling it, alternatively with large blocks, gravels rich in volcanic materials, sands and clays. When the riverbed was totally filled the river started downcutting elsewhere.

More than 2000 faunal remains were discovered in the deposit: straight-tusked elephant (Elephas Antiquus), extinct aurochs, hippopotamus, bear, rhinoceros, horse, wild boar, hyena, wolf, fallow deer, deer, waterfowl, and also fossil leaves of a tree of the Ulmaceae family (Anzidei and Ruffo 1985). The remains of Elephas Antiquus stroke the imagination of people already during excavation, because of their size and quantity: some thirty tusks were found, together with molars, skull and basin fragments, as well as some long bones (Fig. 3).

The human presence is also attested: in 1983 the fragment of a human parietal bone was discovered under a tufa block (Manzi, Salvadei, Passarello, 1990); according to the suggested chronology, this is the time when the Neanderthals occupied Europe. Furthermore, more than 1,500 stone tools were collected, made from small pebbles, as it is usual in this coastal part of central Italy (Fig. 4) (Anzidei and Gioia 1990).

2. Preserving and Valuing the Site

The value of the deposit relies on many issues:

- it is exceptionally well preserved, in the heart of the modern city.
- it allows to reconstruct and depict an ancient landscape as well as to identify the animal species present at the time.
- it shows the presence of hunter-gatherer human groups in the territory of Rome since early times.
- it is the only Pleistocene site to be preserved and visible today, out of the many once existing in the lower Aniene valley.

For this reasons the deposit has always sparked off a great interest among scholars and common people, thus supporting a process of conservation and promotion of the site. The partnership among many institutions made possible the realization of this unusual museum.

The archaeological excavation of the 80s brought to light an area of 1200 sqm. After that, 300 sqm were preserved under a shield of foamed clay, plaster and wood planks. In the following years, up to 1995, the site was totally neglected. In 1996, the site management was assigned to the Municipality of Rome (Sovrintendenza ai Beni Culturali).

The first project for a permanent cover, laid out in the 80s, was aimed mainly at protecting the deposit from natural hazards and...
from vandalism and involved a large building (Anzidei and Morganti 1988).

In the 90s the project was modified and transformed into a museum on the site (Gioia 2004; 2005), adding exhibition areas without architectural barriers. In the year 2000 the new building was completed, and a transparent roof was added over the visitors path. The deposit was then restored in 2001, in collaboration with the Istituto Centrale per il Restauro, and the finds were replaced in their original location.

Museological and museographic actions followed to these works, based on a communication systems that uses textual, verbal, symbolic and technological communication. Thus, in addition to the traditional forms of exhibition, more direct communication tools were preferred, like those of visual and/or interactive type (Antinucci 2004), as well as new educational techniques such as real and virtual reconstructions and hands-on labs. The aim is to make the visitor interact with information and also to keep his/her attention alive.

A great effort was made to communicate not only with an informed public, able to understand on its own the message provided by the exhibition, but above all with the common visitors, who would have more difficulty to understand the contents of the museum (Gioia 2005; 2011). Thus, specific reconstructions were purposefully laid out for the showcases, provided with many pictures accompanied by simple and concise texts. The educational boards use bright colors (each color features a topic) and new landscape reconstructions are also worked out (Fig. 5).

The outside visibility of museum is provided by two large panels of majolica tiles painted with artistic reconstructions of the Pleistocene environment (Fig. 6). The next steps, which will be realized soon, consist in setting up the large outdoor space with a thematic garden exemplifying part of the Pleistocene flora. Here new spaces equipped for hands-on labs will also be created, to carry out future labs of archaeological excavation, stone flaking, use of flint tools, etc. (Bellintani and Moser 2003).

The network of relationships with the territorial, social and cultural background is essential for the future life of our museum (Muscò 2007). Many ongoing activities are designed and planned inside the framework of the close relationships established with the urban study and research centers, mostly the neighboring ones. Thus the museum has set up contacts with the nearby Rebibbia jail, local and city environmental associations, primary school pupils and large groups of university students who carry out their internships there.

3. The Digital Project

The exposition of the many different dissemination features developed by the Museum’s team through time, are aimed to
show that the Museum itself is, since its origin, focussed on the need of a narrative approach, in broad sense and towards very different user targets. The meeting with digital technologies was then particularly easy and effective, as the system developed by the Institute of Technologies Applied to Cultural Heritage of Italian CNR, is going to enhance such a dimension, by an immersive reconstruction of a spot of Pleistocene everyday life.

Starting from the already quoted detailed studies on the geomorphological and environmental conditions, the virtual reconstruction process led to a highly impressive application, through the most advanced photo-realistic terrain generator software and fluid simulators. The whole communication system implies 4 applications:

- The current reality: The first application will create a spot-lighting system allowing a “flash-telling” of the archaeologic deposit’s content. A female voice (actress) will tell what you see today.

- The Plei-sto-station: After the progressive immersion into the Pleistocene world, in the exhibit hall, which allows to see closely the findings and better explore some themes, visitors can also play with “plei-sto-station”. The Plei-sto-station is an educational application targeted to childhood’s learning, implemented by a touch-screen interaction dynamic (Fig. 7). The goal of the four quests is to find the every-day life elements in common between the upper (Pleistocene kid) and the lower (modern kid) domain, thus checking couples of objects referring to the same functions (fruits, leaves, elephant teeth, etc.)

- The flood simulation: One couple of projectors will be targeted to the deposit, and aimed to show on the Pleistocene ground, in the darkness and with an immersive audio track, the effect of river refilling (Fig. 8 a, b).
4. 3D Reconstructions and Filmic Rendering

The 3D reconstruction of Aniene river Pleistocene landscape has been carried out in about 2 working years, with a careful study of the plant and animal situation, and a strong interaction among modelers, paleontologists, geologists and palaeo-botanists.

Such a pipeline led to the realization of

- Photographic relief of the area.
- hi-res modelling of the whole river-bed through cloud-computing computer vision.
- Optimized model and texturing of the 3d mesh.
- CAD 3d modelling of the museum through the Rome City Authority maps.
- Fusion of the two models (Dem riverbed and museum floor for the flood simulation).

For the reconstruction of the ancient valley geomorphology, apart from the geological map, the contour lines at an equidistance of 10 meters were used, as well as the known local distribution of the paleolithic sites.

The following considerations lie at the base of the geomorphological reconstructing hypothesis:

- (evidently) the river system must have passed in every location where deposits of the Saccopastore formation are preserved;
- all the surface elements now below the top level of the site (+34 m a.s.l.) must be the result of later fluvial incision; this is a simplification presuming that – at the time – the valley level was at about 34 m a.s.l. or slightly below; it is virtually impossible to exactly reconstruct the original morphology of these (disappeared) landscape elements;
- where older formations crop out at a level above 34 m a.s.l. for sure a valley has not passed during the Upper Pleistocene;
- when there is a steep side slope far away from the present river, we may presume it may have been formed by the palaeo-Aniene System; if a steep slope is located close to the present river, we cannot guess when it was eroded.

On the basis of such considerations the hypothetical path of valley was drawn. For an area comprising the whole lower Aniene basin, a similar detailed procedure was not possible. Fortunately, it’s highly probable that in the upstream area the river formerly had not steered away much from it’s present course, being more or less deeply incised. This is confirmed by the total absence of the Aurelia, Vitinia and Saccopastore fluvial terraces.

When the geomorphology of the region was satisfactorily shaped, the digitalized model of the riverbed obtained through cloud-computing computer vision was cast into it, in order to reach a whole model as basis for the graphic reconstruction of the Pleistocene landscape of the region, to be used both for the movie settings and for the flood simulation.

The remaining work consisted in the landscape generation through the plant library arrangement (each plant variety was defined and shaped according to scholar’s guidelines, in the mainstream of the detailed paleo-environmental analysis exposed above).
Flood simulation has been performed through specific 3d software based on fluid dynamics: HYBRIDO (HYBrid larRge dImension Liquid sOlver), as waves/splashes/foam simulator by the volumetric grid “Grid Fluid Domain”.

The first step is the parameter starting setting for the algorithm (number of particles for second, longevity, starting dynamic impulse, gravity, fluid viscosity, etc.). Then, the model geometries are imported. The particles use them as rigid bodies to impact on. At the end of each frame calculus a cache is saved, containing informations on the dynamic evolution of spatial coordinates of each single particle (about 434 Gb folder).

For particle foam renderings (Fig. 10), specific plug-ins have been used. Data on foam particles dynamic have also been generated by the Grid fluid Domain on the basis of informations already stored during the calculus. The software allows to directly render point clouds generated, with no need of mesh creation as for the fluid. Afterwards, by a highly customizable flux diagram, it is generated the algorithm to assign a color gradient according to particle speed, vorticity, etc.

The particles, rendered in a picture sequence (TGA 32 bit: 24 bits RGB and 8 bits extra for alpha channel), may be superimposed in post-production to the fluid pictures already rendered by the Global Illumination engines, obtaining the final sequence.

The Terrain general model has been created from 3 DEMs, elaborated inside the GIS software, exported as grey-scale geotiffs, and modeled by displacement mapping technique. The general largest and low-res DEM is 50x40 km and is used to define the background skyline (Colli Albani Volcano etc.). Over that, a more detailed one has been placed (1x1 km), then, the ancient riverbed mesh was combined. All DEMs are linked each other by geocoding, such a careful correspondence between real and virtual landscape allowed virtual camera transitions from viewpoints corresponding to visitors position: the museum footbridge.

Then, DEMs have been textured with procedural plants and materials, starting both from basic software libraries and customized ones, according to pleistocene species, (seasonal) aspects, height, density, distribution by closeness to water streams, soil composition, etc. Such a process was developed in deep relationships with archaeologists, biologists and geologists, who determined even shore and stones shape.

The Elephas Antiquus character has been created from the model of the elephant teeth in the Casal de’ Pazzi Museum, obtained through photographic relief and data post-processing through cloud-computing computer vision, which resulted in a hi-poly model. It was then optimized through polygon re-topology, roto-translated and put in the right scale inside the 3d elephant character (the skeleton and the appearance have been validated by paleontologists). Afterwards, character rigging and animation were performed.

The most part of the film rendering was performed through biased rendering engines: the plugins communicated with 2 external software; for the landscape, wind, atmospheres, and for the interaction between the character
and the river water. In the scene of drowning, the elephant floundering in the water produced a lot of waves and jets, whose calculus had been previously executed and the continuously changing mesh on the river surface was imported by the plug-in frame by frame.

5. Conclusions

The Casal de’ Pazzi Museum narrative complex, and its applications have been presented as a pipeline working model (from the paleo-environmental reconstruction to the public dissemination).

A general overview of the virtual museum applications shows a faceted and complex situation: one the one hand there are successful and long lasting examples, on the other, a huge number of works which disappeared at all after a few years (Pescarin in press).

The Casal de’ Pazzi Museum represents an important case study because of its peculiarity. We have pointed out the singularity of the site in terms of its archaeological features. There are as well some relevant issues as a virtual museum case study, as – to stress the value of a narrative approach instead of the emphasis on technology – it seems particularly meaningful the case of a museum referred to the ancient prehistory, without architectural elements, whose ability to tell its own story is low and where the communication is highly challenging. For such reasons the Casal de’ Pazzi Museum will be monitored and evaluated in its future development. At the moment, preliminary data on the educational impact on schoolchildren (Pagano 2012) are very interesting and promising.

References


Back into Pleistocene Waters
Augusto Palombini et al.


Etruscanning 3D: an Innovative Project about Etruscans

Eva Pietroni
CNR Institute of Technologies Applied to Cultural Heritage, Italy

Daniel Pletinckx
Visual Dimension, Belgium

Wim Hupperetz
Allard Pierson Museum, The Netherlands

Claudio Rufa
E.V.O.CA., Italy

Abstract:
Museums have a lot to gain in terms of communicative impact with new narrative scenarios that are developing through advances in technology. We will discuss the 3D reconstruction of the Regolini Galassi Tomb, a Virtual Museum developed in the Etruscanning 3D European project framework, which engages users with natural interaction interfaces to explore the virtual reconstruction of this famous Etruscan tomb. The development revealed two advantages. Firstly, new perspectives were gained by the archaeologists studying the tomb through the use of 3D reconstructions, visualizations and modelling. Secondly, by installing the application as part of an exhibition on Etruscan culture, visitors were given the opportunity to interact and immerse themselves in the virtual environment using natural interaction interfaces. The application was tested and evaluated in the Virtual Museum Laboratory in the Allard Pierson Museum, from which it is clear that this case study is a powerful example of how important the use of 3D reconstructions is both as a research tool and immersive visualization.

Keywords: 3D Reconstructions, Virtual Reality, Natural Embodiment, Learning, Museum

1. Introduction

Etruscanning 3D is a two-year project founded by the European Commission in 2011 within the Culture 2007 framework. It focuses on the investigation of new digital visualization techniques, to re-create and restore the original context of Etruscan tombs. The project aims – through the cooperation of museums and research organizations from 3 European countries – at the digital acquisition, digital restoration and 3D reconstruction of Etruscan tombs and object collections and their presentation to the public through innovative VR systems. These presentations are proposed at exhibitions in the Netherlands (two important exhibitions have been already open in Amsterdam and in Leiden), Belgium and Germany and for permanent use in Italian museums. Two important Etruscan tombs have been taken in consideration: the Regolini Galassi tomb, in the Sorbo necropolis in Cerveteri, and n.5 Monte Michele tomb, in Veio. The finds from these tombs are exhibited in museum collections and the existing (empty) tombs are not always open to the public. By making 3D reconstructions of the tombs and of the objects which were originally found inside, we can reintegrate what is illegible, contextualize what is fragmented and isolated, re-creating the ancient contexts and relinking cultural ties essential to the cultural understanding and transmission. So our final goal is the cognitive, perceptual and communicative enhancement of the cultural objects that translates into a wider and deeper exchange with the visitors.

Corresponding author: eva.pietroni@mlib.itabc.cnr.it
A 3D reconstruction is not simply a digital replica of a real tomb: its aims is to create an experience that can bring visitors inside the ancient etruscan mind and culture.

The Regolini Galassi tomb is one of the most remarkable Etruscan tombs dated from the middle of the seventh century B.C., famous not only for its rich contents, but also for the many objects that show the Orientalising influence (Fig. 1). This is the tomb we have reconstructed and implemented in a VR environment whose most innovative aspect is the new paradigm of interaction based on natural interfaces. This choice originates from our previous experiences in the field of virtual reality environment inside museums (especially archaeological museums), revealing that traditional devices of interaction are still quite difficult to manage by a large part of public, not literate with digital technologies. For this reason from a couple of years we are focusing our research on the development of new paradigms of interaction based on natural interfaces. This means that the public has the possibility to explore the 3D space and interact only through the body movements (reality-based interfaces), in the simplest and natural way. This is possible thanks to a simple low cost sensor, a depth camera, that does not require the user to wear any marker and does not need expensive licenses to operate. The availability of such sensors, coming from the video game domain, will likely stimulate the use of natural interaction solutions within museums for years to come. The possibility to collaborate directly with Museums in the Etruscanning framework represents a fantastic opportunity to create a strong and concrete link between research and communication in the field of virtual museums, testing the effective impact in terms of cultural transmission, learning and appreciation both in non-linear narrative plots conception and in novel metaphors of interaction. Using the Virtual Museum Laboratory in the Allard Pierson Museum, different versions of the VR application dedicated to the Regolini Galassi Tomb has been (and will be) tested and evaluated with different user groups and this opportunity is a great support for the improvement of the interaction design and for the best contents implementation. Many installations have been made in the past without sufficient testing or evaluation. Museums, technical partners and researchers need evaluation laboratories in order to better understand the different aspects within the mix of technologies and narrative techniques.

2. Natural Interaction and Embodiment in VR Environments

If the development of technologies has entailed a process of extension of the body to enable us to control the environment more efficiently, it offers the ultimate possibility of the displacement of the material body from the confines of its immediate lived space (Featherstone and Burrows, 1995). The crucial point is the capability, even if still at pioneer level, to establish relations and, therefore, communications between natural and artificial worlds based on perceptive-sensorimotor dynamics, instead of symbolic codes: gestures, proprioception, images, sounds, a reciprocal exchange of signals in time and space, on which the original and much more evolved approach in human learning is based. Natural interaction re-configures radically boundaries between natural and...
artificial subjects, their consistency and the “outside” world. We could say that the fundamental division between technology and nature, artificial and human is beginning to blur (Annunziato and Pierucci 2006).

Inside a virtual reality environment, and moreover using natural interfaces, the user feels spatially embodied in the system; this embodiment constitutes a new frontier of the communication and learning processes. In ecological thinking (Bateson, 1979) learning spreads from the capacity to produce difference between the organisms and the ecosystems. The more the conditions of the ecosystem change in an unpredictable way the more we can establish relations and feel incited to continue the interaction because we acquire and exchange information within the environment and become able to create new mental maps. Therefore is the embodiment able to produce difference? The embodiment is based on the principle of enaction; this term was introduced by Bruner (Bruner, 1968) for designing the knowledge is coming through action and constructed on motor skills. Then the concept was re-elaborated by Francisco Varela according to a neuro-phenomenologic approach based on the assumption that the cognitive activity is “embodied” (Varela, Thompson, Rosch 1991), or not separated from the body perception. In these terms information could be identified, communicated and learnt only within a specific context. It seems evident that embodiment depends on the level of the engagement inside the cyberspace: the communication capacity, the sense of sensorial immersion and presence, the emotional and cognitive involvement, the feeling to find correspondences between the real and virtual dimensions. For some years video-games have developed advanced metaphors, introducing multiuser virtual environments, embodied cyber communities, artificial intelligence and natural interaction. They have made the public used to very involving, immersive and sophisticated scenarios, advanced languages and media. Research in the field of digital transmission in the Cultural Heritage sector is still pioneering and these inputs coming from the cognitive sciences and re-used by video-games can be of great interest. We believe, in fact, that, in the archaeological domain, embodied behaviors implemented in virtual reality systems together with natural interaction interfaces represent a really new opportunity. This is especially the case inside museums as they should enable the transmission and fixing of more knowledge in a shorter time (Forte et al. 2008). In other words, they represent a stimulating and innovative gateway to the simulation and reconstruction of the past, a catalyst of learning (Forte et al. 2002; Forte 2008).

In cyberspace the 3D digital model represents the background; the foreground is created by “floating information”, storytelling, metadata, interpretative layers, tools of visualization that can be now accessed through reality based interfaces. The information is organized and integrated in the 3D space, the user is surrounded by it and he can perform actions in order to gain it. In the traditional VR applications, the avatar is a character whose actions are totally controlled by the user, because it does not have its own behaviours. The use of the avatar, a digital personification of ourselves, increases our sense of presence and embodiment in the virtual world, and allows us to better perceive our movements, the proportions of the space, the effects of our actions and interaction with other beings (Annunziato et al. 2008). Thanks to natural interfaces the traditional concept of “avatar” can be overtaken: now the user is directly immersed and acting in the cyberspace, he doesn’t need any longer to be represented by a digital “medium” because he can directly touch the objects, being recognized by the external world. In all our recent VR applications using low cost and markerless natural interaction systems we have observed and studied the public’s feedback (three projects have been developed until now: 1)The Approval of the Rule: virtual experiences among Giotto’s characters, 2010 (Pietroni and Antinucci 2010); 2) Etruscanning 3D, 2011-2013, 3)Virtual Exploration of the ancient
Pharmacy of S. Maria della Scaletta Hospital at Imola, 2012) (figures 2, 3). The evaluations carried out on these projects has shown that visitors appear very involved in the experience and they remain playing with the objects for several minutes, repeating the actions until they have full control on the system, listening many times to the audio explanations and narratives. Moreover, people reveal a creative/enphatic approach in performing the required movements, as they have the feeling of being in a game, exaggerating their role. For this reason the use of natural interaction can produce less precise input, in comparison with traditional interfaces (mouse click, keyboard, joystick....) but this limit does not generate frustration in the visitors, on the contrary it translates to a challenge encouraging people to try, explore and learn until they obtain good results (Alisi et al. 2005; Castellano et al. 2007). We are working in the definition of a proper grammar of gestures, continuously testing it on public, also with the support of experts in cognitive science (Cantoni et al., 2004). In fact gestures need to be really intuitive, responsive and well designed by the authors as not all people have the same perception, coordination and awareness of their own movements. We believe these three study case can be considered pioneering, as at the moment there are very few projects in the world dedicated to the communication of cultural heritage in museums using natural interaction interfaces in 3D real time environments. Instead, this kind of approach seems to be a bit more diffused in 2D visualization, in order to browse through images and multimedia.

3. Methodology

3.1 Research Tool

The techniques we used for 3D representation of the artifacts and the Regolini Galassi tomb itself were various, they included: laser scanning, fotogrammetry and computer graphics. Each tool was selected according to the typology and topology of the objects. Moreover as the process of virtual reconstruction of the Regolini Galassi tomb tries to visualise it at the moment it was first closed, we have been forced to consider very practical questions regarding the placement of the objects, their original shape, color and position. This was not an easy task, as when the tomb was discovered in 1836, by the priest Regolini and the general Galassi, it was not methodically documented and much of the information on the exact location of the objects within the tomb was lost. The first publication of Regolini Galassi (after the preliminary report of the Grifi and Braun) is due to L. Canina which reproduces the first map with the disposition of the grave goods. The plant made by Canina was also used by
L. Grifi in his book of 1841. Some differences in the location and characterization of the objects can be noticed: different sizes and even shapes from the real for some objects, diverse location in the tomb, the size of the tomb itself is different from actual one.

The reconstruction here proposed is neither complete nor definitive, but it is an effort to find the most reliable solution. It is based on the reinterpretation of the first documentation of the nineteenth century and therefore receives only part of the location of the grave goods finally proposed by L. Pareti in 1947 (Pareti, 1947), (Sannibale, 2008). We had to re-evaluate and re-interpret all of the available, unclear, sources to seek answers to these difficult questions (Fig. 4-5). Thus, the process of creating 3D reconstructions is confronting researchers with new questions and (visual) perspectives that is stimulating the debate on archaeological interpretations. They impact not only the way we look at sites and spaces, offering new depth and insight, but they also develop through a multidisciplinary way of working, since the development of 3D reconstructions requires considerable teamwork.

A good annotation system is going and needs some more implementations in order to make it accessible to public to ensure that the effort that has been put into these 3D models is not lost. At the moment a blog is on going where we document the progressive phases of the work (http://regolinigalassi.wordpress.com/ Virtual Reconstruction of the Regolini Galassi Tomb).

3.2 Digitization

The above mentioned interpretative aspects, as far as the digitization techniques, are better described in the second paper the authors have presented in these Proceedings and that we recommend to read in order to have a complete understanding of the projects and its methodologies: “Etruscanning 3D project. The 3D reconstruction of the Regolini Galassi Tomb as a research tool and a new approach in storytelling”. We give here only a few basic details so as to give a general outline understanding. Several types of data have been acquired and elaborated, according to the typology and topology of the artifacts, including point clouds from laser scanner, photogrammetric data (dense stereo matching), and manual modelling. A “time of flight” laser scanner (Riegl z390i) was used to acquire the tomb in 3D as point clouds with high resolution (6 mm) and maximum accuracy of 2-3 mm. The 3D point clouds were postprocessed to obtain a mesh model textured using the ortho-photomosaic obtained from digital photos. From the high resolution geometries (8 million polygon), normal maps were calculated and applied to a low poly version of the tomb, optimized for the real time engine using natural interaction interface. Moreover, from the 3D model of the tomb as it exists today,
we have expanded upon the model to present the tomb as it could have been in Etruscan age; with the objects contextualized inside, based upon historical sources and archaeological interpretation (fig 6). The objects found within the tomb were digitally acquired at the Vatican Museums; in some cases we could use already existing photos, in other cases we had to capture new images of the objects, taking them out of the showcases. The museum did not allow us to use any other equipment, such as a laser scanner. In several cases we used a turn table in order to take photos all around each object, a white circular tent was also used to diffuse the light in an homogeneous way, putting the object inside and avoiding as far as possible reflections on its metallic surface. This acquisition technique was useful especially for dense stereo-matching techniques (Fig. 7). The third method, manual modelling, has been used in case of objects with an accurate and trustworthy visual documentation (photos, drawings, measures) , or in case of objects for which it was not possible to obtain new images (e.g. because they are in the showcase), or, finally, in case of objects with problems of reflection or with very complex details, difficult to digitize with more automatic techniques . The modeller tried to reconstruct the object according to the available documents, paying attention to the correct dimensions (acquired from literary sources or through new measurements taken by callipers) and to the aesthetic aspects.

The ceremonial golden fibula testifies the good results that can be achieved through the manual modelling.

On the basis of photo interpretation and especially of the iconographic comparisons and similarities, we started the work of digital restoration on the deficient objects and decorations. Drawings in grey scale were elaborated upon in order to use them to generate normal maps and apply them to the 3D model. The editing of the final material and shaders was done in Unity 3D, the real-time engine on which the VR application is also based, as was the final composition of the scene.

4. Virtual Exploration of the Regolini Galassi Tomb: VR and Natural Interfaces

In this paragraph we are going to describe more in detail how the virtual reality application dedicated to the Regolini Galassi tomb in Cerveteri has been conceived and developed. The system has been derived from the new generation of games but, for the first time, it has been applied to VR environments dedicated to archaeology and tested inside museums. A first version of the application was presented in some exhibitions: Riches and Religion of the
Etruscans - Princes and Priests (in Amsterdam) and Princesses and Goddesses (in Leiden), open from the 13th of October 2011 until the 18th of March 2012 (Hupperet et al. 2011), and during the Archeovirtual exhibition in Paestum at the Mediterranean Archaeological Tourism Exchange, in November 2011. The choice of natural interaction interfaces influences not only the metaphors of interaction but also the general design and structure of the applications: interacting by movements is completely different from using a joystick. The expectations of the user changes, so the storytelling, the duration of narratives, the selection mode, the relations that can be established with the objects, the manipulation possibilities need to be thought in a different way.

The Kinect interactive area has a cone shape, 5m long x 5m wide (the device is the vertex of this cone). In effect the application was adapted for a smaller area, (3.5m long x 2 m wide), due to the limited dimensions of the room. The dimension of the projection was 12 square meters.

4.1 First Development

In the first version implemented the only user interface was a carpet laying on the floor with a simplified plan of the Regolini Galassi Tomb, on it some hotspots have been placed, in order to suggest the interactive points to the public (Fig. 8). By moving from one hotspot to another the user makes the virtual camera move in the virtual space of the tomb, with proper animations. Once the position is reached by the camera, it focusses on the surrounding objects: a princess and a warrior, alternate in the explanation of the functions, the symbolic meanings, the cultural provenance and the preciousness of each object (Sannibale, 2008). The camera rotates and gets close to each one, a spotlight turns on and the remaining space becomes darker; at this point the voice emerges (about 1 minute of narrative). Each hotspot has 4-6 associated objects. The storytelling aims at creating an emotional involvement (Antinucci 2004, Ryan 2001): the two characters of the buried personages, the princess and the warrior, tell stories from their own point of view, while remaining aware of the times in which their visitors live, so that they are able to comment on the archaeologists’ doubts about unknown aspects of Etruscan civilization. The user can explore the tomb in every order he wants, he can leave the active hotspot at any moment and in this case the storytelling interrupts; if he comes back to a previous hotspot, the narrative begins from the same point it was interrupted. If the interactive area is left empty for more than 10 seconds the system will reset.

The application is mono-user, even if the presence of some people in the interactive area doesn't cause problems (the last input is followed). While the active user is guiding the system, other visitors can sit down in the viewing area, visualizing and listening to the cultural contents in a passive way; always with the opportunity to engage in active exploration (Fig. 9).

This structure is very simple and easy to understand (something between an interactive movie exploring the story and a VR exploration, controlled by natural interfaces) and it does not stress the public at all. In the meanwhile the user has the feeling of being the main actor on a stage. Another fundamental result is that people of every age, technical skill and conditions, even the physically disabled, can enjoy without any barrier.
The medium time of interaction for each user is about 12 minutes; a very good result according to our expectations. In the initial version of the application, we implemented an algorithm of “path finding” for the camera movements among the hotspots; this made animations every time different because they were not controlled by the user or by the programmer through a pre-defined path but directly by the computer. The problem was that some camera movements were very fast and close to the wall, which for some, could cause a sensation of motion sickness. For these reasons, in a successive version we have given the camera some constraints: it moves forward and backward according to user’s position and makes slow and limited rotations, only in relation to the two lateral cells, controlled by the programmer. The application has been built in Unity 3D for MAC and uses the Kinect sensor for motion capture. The framework to interface between the Kinect and the computer is OpenNI; an open source application programming interface (API) developed by Prime Sense and Willow Garage industries for writing applications using natural interaction. This API covers communication with both low level devices (e.g. vision and audio sensors), as well as high-level middleware solutions (e.g. for visual tracking using computer vision).

The small dimension of the interactive area in relation with the real dimensions of the tomb (about 12 mt long) obliged us to use the hotspot as main interface for the 3D exploration instead of a real walk in scale 1:1 (this could have been the best solution probably but not easy to perform inside museums). The application need no calibration, as the skeleton does not require measurements; in fact the only input is the user’s movement and position on the floor, and no other gestures are used. In any case the last version of OpenNI framework (and also of Microsoft SDK), at the moment of this writing, does not require calibration any longer, even if more gestures are implemented.

4.2 Second Development

Despite the fact this solution appears quite satisfactory from a user perspective, we have continued with further experiments in order to define a more complex approach to natural interaction, improving an evolving grammar of gestures and more levels of freedom. This new improved version has been presented in November 2012 at the Science Festival in Genova, next year for a temporary exhibition in Tongeren (Belgium) and a permanent installation will be hosted in Vatican Museums (Museo Gregoriano Etrusco) where the precious funerary goods coming from the tomb are preserved and shown. This version is no longer based on an interactive map on the floor, but it allows the user to explore the 3D space from a fixed position in front of the projection, using his arms to walk through the ancient tomb, full of objects, and rotate his point-of-view. So moving in the virtual space no longer requires the user to walk in the interactive area in the museum, hence presentation is also possible in a smaller space. In this new configuration, four hotspots on the floor are used for the following functions (Fig. 10):

1. hotspot Languages: choice of language, Italian, English, Dutch;
2. hotspot Exploration: free exploration, using arms for translation and rotation on xyz axes;

3. hotspot Selection and Storytelling: the storytelling will be referred to the objects staying in the area in which the user has stopped. Each object, when selected with the right hand, has a dedicated camera and spotlight; it can be animated while speaking;

4. hotspot Start: a short tutorial is given to the user in order to teach him how to interact inside the 3D space. This section includes also an historical introduction about the tomb.

This solution appears to be the best compromise, suitable for every profile of user: all users can immediately understand how it works without any frustration while retaining an active role and feeling much more embodied in the 3D space. In fact the second and forth hotspots aims at real-time interaction; the third hotspot allows one to relax and listen to the stories. Moreover, despite the need of more gestures, it does not require the user to use legs for specific movements, except for moving from one hotspot to another; this means that also people physically disabled can use and enjoy the interactive space. This second version of the application also uses Unity 3D but it runs on Windows platform so we can use the official version of Kinect with the Microsoft SDK.

4.3 Evaluation

At the moment the evaluation has been done only for the first version of the installation, presented in 2011 in Amsterdam, Leiden and Paestum. On the base of these results we have consolidated some aspects and improved weaknesses in the second incarnation. The approaches to evaluation have been many and varied, based on the intended aspect of the application to be evaluated. For example, an observational approach was used to identify the user demographics and determine some of the technical issues that could be refined in future versions of the application by the developers, while interviews with users provided more detailed information on their experience of interaction and immersion in the application. The evaluation put in evidence that the majority of users found the content was inventive, clearly structured, captivating, enjoyable and the application easy to use and to understand.

When asked to select which element (the Etruscanning installation or the Etruscan exhibition) provided the clearest information about the Etruscans, 47% of users identified the exhibition as the best source for general information on the Etruscans. Alternatively, when asked to select which element presented the clearest contextualization of the objects, 47% of users identified the Etruscanning installation as the best solution.

When participants were asked if the integration of the application into the exhibition corresponded with their idea of what a museum experience should be, 71.4% of participants absolutely agreed. This suggests that museum visitors may consider a “museum experience” to be much more than simply looking at objects. In the second paper on the Etruscanning project presented by the authors in these Proceedings, detailed information about the user experience and the evaluation results are given.
In conclusion we believe that prioritization of museum information has to be based on data contextualization, cultural relations and narrative themes. The more the information enhances connectivity, the more visitors will be able to assimilate and elaborate cultural contents. In this way, the museum becomes a place of experience, emotional and conceptual involvement.

References


1. Introduction

Archaeological museums can be uninteresting to many people because they do not connect to the personal narratives that visitors carry with them and, implicitly or explicitly, constantly re-build. Indeed, memory institutions need to sustain, not to say reinforce, their attractiveness and the interest of their visitors if they do not want to find themselves standing still “on the conveyor belt of history” (Serota 1996). They must make cultural heritage more engaging, especially for the young generations of “digital natives”. A challenge for cultural heritage sites is to capitalise on the pervasive use of such media, while also facing the competition from the leisure-based entertainment industry, which attracts visitors through spectacular exhibits and events with experiential, but also even educational and cultural qualities. However, digital cultural heritage content and assets can be expensive, technically difficult to make, and hard to renew.

This is where novel research kicks in. Recent investigations in interactive digital storytelling, personalization and adaptivity, and mixed reality, coupled with mobility-enabling systems, promise to make not only cultural heritage sites more attractive but also to provide new means to leverage and exploit the existing digital libraries that have been developed since several years in the cultural heritage world.
The aim of our paper is to present the on-going EU funded project CHESS (Cultural Heritage Experiences through Socio-personal interactions and Storytelling - www.chessexperience.eu). The principal objective of CHESS is to research, implement and evaluate an innovative conceptual and technological framework that will enable both the experiencing of personalised interactive stories for visitors of cultural sites, and the authoring of narrative structures by the cultural content experts.

2. A Story of Storytelling

2.1 Storytelling from Oral to Interactive

Storytelling is defined as the production of a narrative that communicates experiences in an oral form (Villaseñor 2007). As any communication process, it is an action that reproduces but also produces culture. This means two things. On the one hand, it contributes to re-experiencing one’s own heritage, thus reinforcing identity and the feeling of belonging to a community (Abrahamson 1998). In this sense, storytelling fulfils also a moral role since, through their underlying messages, stories transmit cultural values and sanction what beliefs and behaviours are allowed or not (Bruner 1990). On the other hand, because storytelling is a general human capacity (Schank 1995) and because it deals with personal emotions and experiences, it helps to overcome cultural distances and to be able to understand other ways of living and thinking (Bruner and Turner 1986).

Storytelling constitutes the expression of an experience but also an experience in itself (Johnsson 2006). This implies face-to-face communication, and therefore an interaction between the storyteller and the audience. Thus, storytelling is per se an interactive performance, in which the teller adjusts the vocalization, wording, physical movements, gestures, and pace of the story to better meet the needs of the responsive audience (Bull and Kajder 2004). Telling a story is not limited to just oral communication, but also involves the creation of an atmosphere through senses, in which all told events are emphasised one after the other (Bailey 1999). In order to tell a story and to catch the audience’s attention, the story teller does not only relate the events, but also makes the audience feel emotions along the narration (Bates 1994). This is where empathy comes into play. Enriching a story, catching the audience’s attention and stimulating the senses contribute to create a moving story, making it memorable (Peterson and McCabe 1984).

Storytelling is considered the first, most essential form of human learning (Bruner 1990). It has been defined as an imaginative form of discourse, which guides the listener to a process of meaning-making in the openness of an “imaginative state” (Bruner 1990). Not only this definition is close to the constructivist theories of learning, but it has been demonstrated that stories are more easily remembered than raw facts because they contain an underlying structure and can be linked with prior experiences (Pozo et al. 1989). Moreover, social theories of learning (Vigotsky 1978) emphasise the importance of interaction between participants. Storytelling complies with it, especially if the audience is encouraged to participate (Villaseñor 2007). Storytelling also stimulates other cognitive factors that contribute to learning (Egan 1989) such as attention (by the correct pace in the story that keeps the listener engaged) and empathy (emotional identification provides a cognitive anchorage, a frame of reference to make sense of and to ground the new information to be added).

The advent of the written word and its systematization with print transformed the oral tradition of storytelling into a literary form of art. The basic rules of oral storytelling are the same: it needs to have a clear purpose, contain strong characters that evolve, depict an atmosphere, and tell a story that is emotionally compelling,
timeless and raises questions (Norris 2011). However, while it has the advantage that it includes images and it can be taken anywhere, it has lost its interactive character and the contact with the human narrator.

The next step in the historical development of storytelling is the cinema. Again, movies apply all the rules of traditional storytelling but enhance the creation of an atmosphere thanks to the audiovisual dimension and introduce new ways to develop the narration beyond Freytag’s narrative model\(^\text{2}\), thanks to the connotative and denotative power of images. Unlike books, movies are not ubiquitous, but they recuperate the social (although passive) aspect of storytelling, since a community (the audience) gathers to watch the film.

During the 90s, the use of computers opened a new field called digital storytelling, which has been defined as the combination of narrative with digital content (images, sound, and video) to create a short movie, typically with a strong emotional component. Because of this multimedia component, digital storytelling allows many of the elements of traditional storytelling to be integrated and to address different learning styles (Springer, Kajder et al. 2004). This is especially true in the case of Interactive Digital Storytelling, which appeared some ten years later with the introduction of interactivity. This new form combines participation, as occurs in computer games, with automatic story generation and narration. User interaction is strongly related with the paradigm adopted to create the story. Plot-based approaches are primarily concerned with the importance of narrative structure, and their main focus is to ensure that a certain level of coherence and dramatic tension is provided, much as it is in more traditional storytelling media. In a plot-based approach (Spierling, Braun et al. 2002), plot generation and visualization are treated separately, and well-defined stages of authoring, planning, and user interference are present (Holmquist et al. 2000).

In character-based approaches (Young 2000), the storyline usually results from the real-time interaction between virtual autonomous agents and the user (Balet et al. 2008). Much previous research has focused on the application of various Artificial Intelligence techniques to imbue automated characters with personality, desires, and goals (Sgouros 1999).

In spite of the several potential advantages of this approach to communicate / experience historical events or other societies (Danks et al. 2007), some authors consider that interactivity and storytelling do not fit together because the user’s actions interfere with the plot and the coherent development of the story is lost, which can be especially negative for highly structured processes such as learning (Crawford 2005).

### 2.2 Storytelling and Archaeology

As member of the historical knowledge domain, Archaeology is deeply related with narration: it does not constitute a goal (as it happened with chronicles) nor it is used as a source (as it happens with History), but it still constitutes the main communication means (Ginzburg 1989). In the tradition of cultural materialism, the product of research corresponded to a chronological, sequential narration about what had happened in broad regions or countries (Trigger 1989). Later, the European processual functionalism and the Nord-American “New Archaeology” attempted to transfer the current scientific paradigm to Archaeology. As a consequence, they established the anthropological explanation of the material remains as their main goal (Trigger 1989), and even attempted to adopt a formal language such as the one used in logic or mathematics.

The postmodernism turn of the 60s, and the consequent theories such as post-colonialism, post-industrialism and feminism, drew attention to the subjectivity involved in

---

\(^2\) Gustav Freytag (1816-1895), German novelist, defined the narrative structure as: exposition – rising action – climax – resolution.
interpretation processes and questioned the meta-narratives that until then had sanctioned the production of scientific, objective knowledge (Villaseñor 2007). For post-modernism there is no such thing as the truth nor universal laws constitute a scientific goal; instead, multiple opinions and perceptions are welcome to understand and represent reality. This opens the way to re-engage with storytelling. Post-modern archaeology recuperated it within the framework of micro-history, understood as subjective interpretation of researchers or as personal stories “told” by those who lived the facts.

Although nowadays a clear distinction between facts and stories, between proof and opinion, still persists in Archaeology, an engaging text that aims to engage or convince the reader needs to have a balance between validity, reliability, and a human or emotional component (Villaseñor 2007). This is even more compelling in an educative context, and is the reason why several Social Sciences choose storytelling as the main way to communicate with non-expert audiences (Nash 1990; Bedford 2001). The specific problem of Archaeology is to find a communication formula that integrates different representation formats (descriptions, maps, pictures...), multicausality, time and space scales, etc. Some authors developed interactive storytelling projects with a multimedia approach (Tringham 2004), but they remained isolated attempts until the spreading of low-cost Virtual Reality and serious games (Roussou 2001; Pietroni, Forte et al. 2006; Kee, Beheshti et al. 2008; Gonzalez-Tennant 2010).

2.3 Storytelling in Museums

Museums are storytellers (Bedford 2001; Johnsson 2006). Starting from the premise that the meaning of archaeological objects is not immediately accessible to non-expert audiences, museums propose one or more interpretations of objects, which are usually presented through a combination of different mediators sequentially located within the space of the gallery. The 19th century museum proposed a monolithic, authoritative perspective, based on a chronological and geographical arrangement, and on labels (Wyman et al. 2011). As a consequence, storytelling was implicit, mainly related to the objects’ historical context, and therefore only accessible to experts. The transformation of the museological practices during the second half of the 20th century (the so-called New Museology) has transformed exhibitions, which now present different points of view (mainly related to the social and cultural context); are based on other arrangements (e.g. thematic); and include different tools for different audience sectors to build their own interpretations or even share authorship with the museum (Fisher et al. 2008).

The adoption of a more explicit storytelling approach to exhibition design contributes to making collections more accessible and engaging for different kinds of audiences: it creates a relaxed environment that raises self-confidence (Johnsson 2006); establishes a universal way of communication; and because it invites the audience to fill in the blanks with their own experiences, it helps to set emotional connections, which are deeper than intellectual understanding (Bedford 2001; Springer et al. 2004). Moreover, traditional storytelling has recently been introduced by museums as another means to contextualise objects in two ways: as specific events in the museum’s family or school programmes (e.g. National Archaeological Museum of Athens), and by including in exhibitions personal stories related to historical events (e.g. the Imperial War Museum North). This has three consequences (Fisher et al. 2008; Hooper-Greenhill 1999): firstly, objects become closer and more relevant for visitors; secondly, social minorities see themselves represented in the museum; finally, it helps the majority culture to overcome the self-centred, reductionist perspective of the world.
At present, museums are trying to develop their exhibitions so that the experience adapts to the different interests and needs of visitors, with the ideal goal to make it fully personalised and interactive. They are also aware of the need to have more powerful tools to support their function as creators of narrative experiences. The challenge of supporting interactive storytelling in Cultural Heritage has been at the core of several previous projects, such as Art-E-fact (Sperling and Iurgel 2003), NICE (Bernsen and Dybkjr 2005), INSTEP (Danks et al. 2007), Brighton Fishing Museum (Danks 2008), or INSCAPE (www.inscrapers.com). However, while the results of interactive digital storytelling projects have been successful when applied to games, films, or multimedia, their potential to enhance storytelling in museums is less clear. Museums comprise a distinctive setting for storytelling, not only because they need to integrate digital media with physical artefacts and settings (Benford et al. 2001), but also because they are very often experienced by groups.

3. The CHESS Project

CHESS proposes to enrich the museum visit through personalised interactive storytelling experiences on two axes: a) by personalizing and (dynamically) adapting information about cultural artefacts to each individual or group of individual visitors, and b) by (re) injecting the sense of discovery and wonder in the visitor's experience. The driving force of the project is its experience-oriented, user-centred approach, which aims at ensuring that its users’ needs are perfectly addressed, thus maximising the acceptance of a highly innovative system and its potential for use in pragmatic situations. CHESS targets two levels of end-users:

Visitors are people experiencing an interactive story created with the CHESS authoring tool. They are invited to join in the available adventures when entering the museum or from home. When on-site, they participate through their mobile phone, receiving information from the system according to the plot, their position, their personal profile, but also contributing information in response to the system’s solicitations.

Authors are non computer-experts (e.g., content providers, curators, and museum staff) in charge of creating cultural interactive experiences for visitors. They use the CHESS authoring tool to create narrative structures that use existing digital content, support several devices and multiple visitors, and adapt to the visitors’ profile, progress within the experience, and interactions.

To support this approach, a user-centred design philosophy is followed throughout the entire course of the project, both in the design and the evaluation phases. The main tenets include:

An iterative process of design – development – evaluation, which begins with a comprehensive analysis of the needs, wants, and limitations of the end-users. For every step of the project, a multi-tiered evaluation methodology has been set, in order to test the validity of the design, either in real world experiments or through the organisation of user workshops.

A participatory design methodology, implemented with a small group of end-users (both museum curators and representative groups of visitors) who, either as partners in the consortium or through a user group actively participate in the planning and design of the scenarios from the outset.

The development of both a personalised and an adaptive system, which delivers personalised narrative experiences for each visitor. The modelling of a visitor profile starts with its assignment to a specific pre-defined persona thanks to techniques such as the

3 The concept of “persona” comes from the Human-Computer Interaction field and corresponds to detailed descriptions of
CHESS Visitor Survey (a brief questionnaire aimed at gathering information from the visitor). Additionally, CHESS may make use of “automatic” extraction of users’ profiles from social networking sites. During the visit, the system will be able to recognise the change in interests and needs over time and dynamically adapt to the visitor throughout the entire visit.

The CHESS consortium comprises seven organisations from four different countries, which provide all the necessary competencies throughout three complementary categories of partners (Table 1).

The different nature of the cultural partners (an archaeological museum and a science centre) provides an interesting test bed for the implementation of interactive digital storytelling in different contexts. Cité de l’espace is a science centre displaying educational models with a high degree of interaction, and it expects that CHESS provides a coherent link between exhibits. The Acropolis Museum (AM), on the other hand, displays originals aimed at contemplation, with a low degree of interaction, and expects from CHESS an explicit interpretation of objects.

### Table 1. The CHESS Consortium

<table>
<thead>
<tr>
<th>Name</th>
<th>Country</th>
<th>Type</th>
<th>Field</th>
<th>Main role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diginext (DXT)</td>
<td>France</td>
<td>Industrial</td>
<td>Simulation and VR</td>
<td>Coordination, Authoring Tool &amp; Distributed Framework</td>
</tr>
<tr>
<td>University of Athens (UoA)</td>
<td>Greece</td>
<td>Research</td>
<td>Computer Science</td>
<td>Story Model, Personalization &amp; Adaptation, User-centred design and scenarios</td>
</tr>
<tr>
<td>University of Nottingham (UNOTT)</td>
<td>UK</td>
<td>Research</td>
<td>HCI</td>
<td>Evaluation</td>
</tr>
<tr>
<td>Fraunhofer Institute of Technology (IGD)</td>
<td>Germany</td>
<td>Research</td>
<td>Visual Computing</td>
<td>Experiencing systems</td>
</tr>
<tr>
<td>Real Fusio (RF)</td>
<td>France</td>
<td>Industrial</td>
<td>Interactive visualisation</td>
<td>Dissemination, Multimedia Assets automated simplification</td>
</tr>
<tr>
<td>Acropolis Museum (AM)</td>
<td>Greece</td>
<td>Cultural</td>
<td>Archaeological Museum</td>
<td>Content development</td>
</tr>
<tr>
<td>Cité de l’espace (CITE)</td>
<td>France</td>
<td>Cultural</td>
<td>Science Museum</td>
<td>Content development</td>
</tr>
</tbody>
</table>

CHESS Visitor Survey (a brief questionnaire aimed at gathering information from the visitor). Additionally, CHESS may make use of “automatic” extraction of users’ profiles from social networking sites. During the visit, the system will be able to recognise the change in interests and needs over time and dynamically adapt to the visitor throughout the entire visit.

The CHESS consortium comprises seven organisations from four different countries, which provide all the necessary competencies throughout three complementary categories of partners (Table 1).

The different nature of the cultural partners (an archaeological museum and a science centre) provides an interesting test bed for the implementation of interactive digital storytelling in different contexts. Cité de l’espace is a science centre displaying educational models with a high degree of interaction, and it expects that CHESS provides a coherent link between exhibits. The Acropolis Museum (AM), on the other hand, displays originals aimed at contemplation, with a low degree of interaction, and expects from CHESS an explicit interpretation of objects.

4. **Development of the CHESS Project at the Acropolis Museum**

The Acropolis Museum (http://www.theacropolismuseum.gr) is an archaeological museum devoted to the archaeological findings of the Acropolis of Athens, from the Greek Bronze Age to Roman times. The first museum was built in 1865 at the top of the Acropolis hill but, in spite of several enlargements (1888, 1946-47), it continued to be insufficient to accommodate the findings from the excavations and, later, the increasing amount of visitors.

The new Acropolis Museum (figure 1), built by architects Bernard Tschumi and Michael Photiadis, opened to the public on June 21, 2009. It is located 280 meters, as the crow
flies, from the Parthenon, by the south-eastern slope of the Acropolis hill. The building lies on the archaeological site of Makrygianni, which contains architectural remains from the Roman and early Byzantine Athens. Today, the new Acropolis Museum has a total area of 25,000 m², and nearly 4,000 objects are exhibited over an area of 14,000 m², ten times more than that of the old museum on the Acropolis hill.

Thanks to the architectural design and to the presence of Archaeologists-Hosts, the museum tells a spatial, chronological, and artistic story. The journey starts at the slopes of the hill and its sanctuaries (on the ascending wide glass-floored gallery after the ground floor lobby) and arrives at the Parthenon (on the second floor), through the Archaic Gallery (on the first floor), where visitors can wander amongst the architectural and sculptural remains of the period spanning from the 7th century B.C. to the Persian Wars (480/79 BC). The flexibility of its museographical design, as well as the diversity of historical facts and approaches behind the objects, makes the Archaic Gallery the perfect context to develop the CHESS project (Fig. 2).

The project started with the definition of the end-user requirements. This was achieved thanks to an ethnographic research (involving observation and interviews with visitors and the museum staff), which provided the basis for two tasks. Firstly, the initial definition of personas: 26 relevant variables were established, from which 6 archetypical profiles (5 for visitors and 1 for authors) were distilled (Fig. 3). The second task was the creation of a scenario describing the four phases of a visit (preparation, arrival, tour of the Archaic Gallery, subsequent activities) by the persona “Natalie Schmidt”. This helped understand the use of the system in context and highlight the critical design issues. In the Human-Computer Interaction field, scenarios are later broken down into a more detailed, visual, step-by-step description of the visiting experience. In the case of CHESS, these storyboards are based on the notion of trajectories (Benford et al. 2009). The Acropolis Museum offers a particular challenge for the implementation of the CHESS system, which is the harmonious integration of high-tech interaction with the contemplation of archaeological originals.

The proposed storytelling model for CHESS attempts to capture and structure the knowledge behind the way humans are creating the stories, and how these stories may eventually be presented to the visitors. Five major phases have been identified in this process, corresponding to the making of a movie or the setting up of a play (Fig. 4).

---

4 Archaeologists-Hosts are members of the museum staff, who amongst other tasks, are available daily at the exhibition areas to answer visitors’ questions about the exhibits.

5 Scenarios are commonly used in the design of interactive computer systems and correspond to informal narratives that describe human activities or tasks and that are intended to provide some structure to guide the work of design. Since they are often presented from the perspective of representative users of a system, they can be linked with the concept of persona.

6 Trajectories are a series of diagrams representing interleaved paths, each of which expresses an individual’s journey through an experience that may extend over multiple spaces, timescales, place participants in various roles, and employ a variety of interfaces.
The first step is the definition of the story concepts, which in the case of the AM correspond to all the possible subjects and perspectives related to the central message conveyed by the exhibits of the Archaic Gallery (Fig. 5). This knowledge was organised in five specific chapters, each of which was developed by internal and external researchers, archaeologists who collected and wrote the basic scientific materials (images and texts). Then, a selection of information is extracted by story authors, including a novelist, who collaboratively write the story script while tailoring it to the characteristics of each persona. In this scripting phase, the basic archaeological knowledge is transformed into emerging stories, with characters, plot, climax and roles.

During the staging phase, story authors are supported by museologists to associate script pieces with exhibits, paths and hotspots in the physical space of the exhibition. In the case of a given museum setting, where the exhibition arrangement is purposefully meaningful, the placement of the story into the physical environment is necessarily concurrent with the previous phase. Next, in the editing phase, authors do the actual montage of the story, i.e. they select (or prescribe when not already available) all the multimedia digital resources and applications that will be employed to realise the defined script. Such resources are generated during the producing phase (including audiovisual material, games, quizzes, AR models, and applications) and are imported into the CHESS library by museum experts or they are generated and contributed by specialised teams of producers (multimedia creation teams). A sophisticated Authoring Tool is being designed and implemented for this purpose, to support authors during all the authoring phases described above.

Finally, in the experiencing phase, the authored stories will be adaptively provided to visitors through web-based and mobile experiencing systems, thanks to several cooperating engines (storytelling, profiling, personalization) and services (location tracking, multimedia presentation and interaction), depicted in figure 6 (Vayanou, Ioannidis et al. 2012).

7 The chapters are: 1. Introduction to the Archaic Period; 2. Description of the Archaic Acropolis (virtual reconstruction); 3. Gods and heroes; 4. Animals and monsters; 5. The world of Humans

8 For example, apart from the most popular exhibits, stories always try to include the less explored exhibits in the Archaic Gallery.
5. Innovative Aspects and Potential Advantages

The CHESS project draws from the storytelling tradition: it is an interactive and multimedia application; its contents will be orally told by a narrator (with written version for audio-impaired users); and its development is based in the movie production model.

From an archaeological point of view, CHESS can be classified as a combination of post-modern and traditional approach because either different historical characters or a scientist (depending on the theme) will tell (personal) stories about the past, to help visitors understand the ideals and values of other societies. On the other hand, the combination of hypertext, multimedia and interactivity fits the narrative basis of Archaeology as well as its multiple representation formats, and allows a flexible non-linear communication (including chronological, thematic and object-oriented approaches), which is not possible with traditional (non-virtual) dissemination means.

Previous projects of interactive storytelling in museums (Tselios et al. 2008; Ardito et al. 2011) emphasise the importance of the context and propose a series of dimensions for design and evaluation. These proposals provide general guidelines for one device in a single context. Yet, such experiences are essentially hybrid in structure and involve multiple spaces (real and virtual), interfaces, and user roles (participants, spectators, bystanders). This is why CHESS proposes the use of trajectories as the most comprehensive and coherent approach and has been developing this former analytical tool through its adaptation, extension and implementation in cultural heritage environments as an early stage designing tool.

In this respect, what is currently missing in the cultural heritage field is a set of tools that assist different authors to bridge the gap between scientific databases and educational interpretations for audiences. Several ICT tools (e.g. Matthew, ARCO or INSCAPE) have been developed but they present several limitations: while some are constrained to the simple exhibition of a particular set of objects and do not allow building narrative-driven interpretations, in other cases, a certain level of programming skills is required. The CHESS Authoring Tool not only addresses these issues but also integrates two other central aspects: on the one hand, it takes into account the specific needs and constraints of the different authors (through the participatory design method); and on the other hand, it includes the users’ profiles and trajectories during the authoring phase.

The personalisation methods currently used in the Cultural Heritage field require that users explicitly give feedback by specifying keywords, providing ratings or answering questions about their interests. In the case of museum visiting, the smoothness of the experience is undermined by the need to fill in data about one’s self. This is why CHESS aims at developing a dynamic system, which will adapt to the visitor throughout the entire visit and beyond. This is done thanks to the combination of several profiling methods, namely an engaging mini-interview that may be completed before the visit; the extraction of users’ preferences and interests from their existing social networking profiles (if they
consent); and finally, implicit methods of personalization (i.e. behaviour during the visit), to suitably and dynamically adapt the visit.

With regard to the visit experience, some museums contemplate the possibility that visitors prepare their visit in advance at the museum’s website or send to their email address the tasks completed at the exhibition’s kiosks (e.g. at the Victoria & Albert Museum). CHESS moves beyond these fragmentary activities by seamlessly extending the visit experience in time thanks to the production of personalised narrative-driven cultural “adventures”. On the one hand, visitors will be able to start their personalised visit from home and include the surrounding archaeological sites. On the other hand, visitors will be able to easily record and upload the most interesting moments of their visit, so that they can continue an individual or “shared” exploration through different social networks.

From the point of view of learning, personalisation and real time adaptivity should benefit learning by tailoring contents and formats to visitors’ interests and skills, and by adapting the experience to the visit conditions. Seemingly, it has been demonstrated in other studies (Economou and Pujol 2007) that interactivity (control over the system, activities) fosters engagement, memorization and restructuration of previous concepts. On the other hand, the indoor location capacity and the Augmented Reality experiencing system (fields in which CHESS will contribute by testing a scalable localisation system using hybridisation techniques) are aimed at enhancing the direct interpretation and exploration of exhibits by visitors, with the possibility to see coloured reconstructions and contextual information on the spot (Fig. 7). Also, the possibility of experiencing different stories is likely to stimulate re-visiting. Finally, the creation of an extended experience (with pre and post-visit online activities) should increase learning, thanks to the previous establishment of conceptual anchors (Pozo et al. 1989) and the subsequent reinforcement of the message.

6. Future Work

In order to successfully achieve the aforementioned innovations and to verify the potential advantages of the application for archaeological museums, CHESS has developed a comprehensive evaluation framework, which can be generalised for use with novel digital cultural storytelling experiences at large. Originally the concept of a museum was “museum centric”. In the last decade, the focus seems to have shifted to “visitor centric”. But neither one nor the other holds; visitors are not empty vessels just like museums are not empty buildings. Therefore, the balanced museum recognises that museums and visitors equally construct the museum experience. Evaluation can be used as a process and a tool to explore the balanced museum’s programs, especially when these involve innovative digital technologies.
CHESS will use a “User experience (UX) evaluation” approach, which allows assessing the quality of the experience, that is, whether visitors engage with the digital cultural experience through the high-level interactive narrative activities performed as intended by the designers and cultural content authors. It is non-trivial to evaluate user experience holistically and come up with solid results, since user experience is subjective, context-dependent and dynamic over time (Law et al. 2009). For example, in the case of CHESS, studying user experience must examine a plethora of parameters in addition to the user’s profile and demographics, interests, location, visit situation etc. These parameters are high-level constructs of user experience that can be used as the basis for studying it and may include, for example, affective response, immersion, cognitive or conceptual change, perception of value, and inspiration. The effect of particular technological choices (e.g., adaptivity, transparent “user modelling” methods via social networks, the mobile Augmented Reality features, etc.) will also be examined in this task. This kind of evaluation can only be carried out with real users in the context of the environments where the experiences are to take place.

7. Conclusions

CHESS proposes to create narrative-driven cultural “adventures” through hybrid structures, which adapt continuously to visitors, extend over space and time, and involve users in multiple roles and with different interfaces. To achieve this, CHESS integrates interdisciplinary research in personalization and adaptivity, digital storytelling, interaction methodologies, and narrative-oriented mobile and mixed reality technologies, with a sound theoretical basis in the museological, cognitive, learning, and leisure sciences. This tightly integrated framework is applied and tested with two renowned cultural sites, the New Acropolis Museum in Athens (Greece) and the Cité de l’espace in Toulouse (France).

With regards to previous projects in the field of digital interactive storytelling for cultural heritage settings, the CHESS project introduces several novelties (adoption of a user-centred approach, personalization through the use of personas, real time adaptivity through the use of localization systems, use of trajectories in the design phase, extension of the visit in space and time). We expect to demonstrate that these will positively contribute to enhance the dissemination of archaeology in museums.

Acknowledgements

This work has been supported by the CHESS Project (www.chessexperience.eu), which is a three-year Specific Targeted Research Project (STREP) co-funded by the EU’s 7th Framework Programme under contract number 270198. CHESS addresses the target outcome “d) Adaptive cultural experiences” of “Objective ICT-2009.4.1: Digital Libraries and Digital Preservation”.

We would like to thank all the members of the CHESS consortium who contribute actively to the development of the CHESS project, namely: Steve Benford (UNOTT), Christophe Chaffardon (CITE), Athena Chioti (AM), Nikos Desypris (AM), Katerina Diamantidou (AM), Niki Dolli (AM), Jerome Duchon (DXT), Stamatia Eleftheratou (AM), Timo Engelke (IGD), Monique Geyres (DXT), Jesse Himmelstein (DXT), Raphael Jacob (AM), Leo Kaloyrnas (AM), Akrivi Katifori (UoA), Jens Keil (IGD), Boriana Koleva (UNOTT), Natalia Manola (UoA), Panagiota Marandidou (AM), Alexandra Nikiforidou (AM), Prof. Dimitrios Pandermalis (AM), Olivier Prat (RF), Stefan Rennick-Egglestone (UNOTT), Holger Schnadelbach (UNOTT), Hara Sfyri (AM-RF), Lydia Themeli (AM), Peter Tolmie (UNOTT), Ky-Nam Tran (RF), Manolis Tsangaris (UoA), Ilaria Valoti (CITE), and Christina Vlassopoulou (AM).
References


Abstract:
This project consists in the development of an installation involving several technological elements which complement each other aiming to provide tools for interpretation and experimentation. Among other technologies, the installation uses Augmented Reality, 3D Modelling, Projection Mapping, and 3D Printing. Part of the archaeological site of Portus, Italy was presented as an example for the visualization, focusing primarily on one of the warehouses. The methodology developed here is transposable to many archaeological sites, and its main goal was to show the mentioned technology at work, which may be used to assist in the interpretation of the site. As we progressively become more multi-skilled, the boundaries between industries and disciplines become less visible and it is difficult to pre-plan when it will be required to bring other specialists into the field that is to be developed.

Keywords:
3D Printing, Augmented Reality, Museum Display, Interdisciplinary Research.

1. Introduction

On the one hand, archaeologists study the relationships between human behaviour and material culture (Skibo, Walker, and Nielsen 1995). On the other hand, Psychologists have studied the relationship of how the audience engages and retains the information given in a museum (Shapiro and Kemp 1990). Multimedia and new technologies allows users to engage and retain this information through a different pathway. Creating a space for information and exchange, the mix of technology and historic information can enhance the interest and information retention of the audience. This project attempts to demonstrate this through experimentation with a museum element object, looking also to emphasise the benefits of multidisciplinary and multi-industry collaboration.


Through visual communication methods, it is thought that the information can be distributed to a wider audience. As case study, the project focused on the warehouse no.17 of the archaeological site of Portus, Italy.

2. Design Interpretation: the Case of Portus

Portus is a Roman archaeological site in Italy. The Portus Project covers a wide range of archaeological research techniques such as geophysical survey, excavation and archaeological computing methods. This project has been developed in various stages and it has involved the work of several international scholars, presenting a very good opportunity for the implementation of our project.

The Portus Project is comprised of research schemes by the University of Southampton and Cambridge University, and it aims to make the archaeological site a “learning” site. Due to the amount of different
aims and sub-projects undertaken within the main project, a number of working teams have been assembled from several disciplines. The research group promotes the use of computer graphics to assist with the interpretation and to present the archaeological information. The purpose of this has been to enhance the general pedagogic experience and to accelerate the learning curve required to fully understand the site. However, one of the essential challenges of the archaeological site is that there is no evidence of how the buildings used to look in the past. The only information available is found in architectural plans of the building remains, excavation data and the geophysical surveys (Keay 2009). This lack of information presents an adequate case to experiment with the interpretation process for the models.

3. Methodology

The interpretation of the archaeological site has to be based on archaeological and historical evidence. In the case of this project, it was considered the role of the designer to blend these two sources into a visible installation that helped in the visualization of archaeologist’s interpretations of the site. The information for the visualization was based on the Roman site of Ostia, following the work of Diana Kleiner and the previous visualisations from ‘The Portus Project’ of the University of Southampton (Kleiner 2009; Keay 2005). Archaeological, architectonical and historical research along with the views of the project specialists are important part of the interpretive and creative process for the designer, since it is here where an interdisciplinary approach can help to deliver the information.

4. Reconstruction

The reconstruction process in 3DCAD techniques was based on research produced by archaeologists and historians. There is always a chance to bring the artistic element. Archaeologists and researchers can produce a database of textures annotated with metadata in order to document the implementation of such texture on the model. This will provide information to the designer, curators, the team and other groups about their recommended usage. Furthermore, when comparing proposed textures, other users can be able to contextualise such visualisation, documenting their implementation in the object.

The basic methodology used for the reconstruction of the building was as follows:

- Historical research of the archaeological site.

- Study of alternative similar archaeological sites.

- Generic understanding of Roman ancient architecture.

- Image and texture gathering.

Portus presented a unique case for interpretation debate due to the lack of information about the original shape and height of the buildings (Keay 2009). This has led the research team to gather information from other cities and ports within the Roman
Empire. Including elements from external sites invites to debate different visualisation from different research groups. This project provides an affordable environment for such debate and collaboration.

5. Elemental Shapes

The work of Kleiner (2009) covers essential information about the architectural visualisation of sites like Ostia and in the case of this project, the information available alluding to elemental shapes of the buildings and the site were taken mainly from this author.

Due to the predominant influence of Ostia on the history and development of Portus, the former site can be used as reference for the reconstruction of the architecture of Portus. In this case, the visualisation made reference to available evidence from this city. Most of the scholars involved with the visualisation of Portus had retrieved and adopted elements not only from Ostia, but also from Rome and Pompeii (Keay 2009).

Buildings at Portus were constructed on a large scale by order of Emperor Adriano. This is evident in the warehouses and domestic buildings in the traditional cities previously mentioned. The large size of the buildings was partly influenced by Greek architecture (Kleiner 2009). In Ostia, the warehouses (horrea) were used to store all supplies coming into the city and they presented the following two basic designs: internal courtyards with groups of clustered rooms around them or rooms placed back to back in double rows (Aldrete 2004). There was the need to provide housing for the people arriving at the port, so although the warehouses had a pre-designed purpose, there is evidence of their construction in an insular manner (Kleiner 2009; Aldrete 2004). The insular buildings were usually three or more storeys tall in which housing was provided for people with fewer resources (Kleiner 2009). Most of the insular buildings that had any rooms facing towards the street were used as taverns, thermopoliums or shops (Aldrete 2004).

Most Roman buildings had gable roofs and the use of tiles is very common so warehouses and insular buildings are no exception (Kleiner 2009). The use of gable roofs was also implemented in Ostia, Rome and Pompeii and several warehouses and temples within Ostia remain standing as evidence of this.

Within the warehouses there are other elements like the dolia. The dolia were four-walled enclosures underneath a surface consolidated by ceramics. They were used to store liquids such as wine or olive oil (Aldrete 2004). In addition, due to the architectural preferences of the Emperors Adriano and Trajan, there are several ornamental elements that are also present within the warehouses. Elements such as pediments, friezes, architraves and plates would form part of the main entrance (Kleiner 2009). They were meant to embellish and imperialise the entrance, and at the same time, the whole building. The use of arches and pillars for the main entrance within Portus is emulated from the main entrance of one of the warehouses from Ostia. Since Roman society paid particular attention to details, items like stucco mouldings and the Dent du Loop are always present within many architectural objects in Ostia, Rome and Pompeii (Aldrete 2004; Kleiner 2009).

By the 3rd century AD, there is evidence that buildings were being re-occupied in Ostia. At the time Portus was adopting the role of an industrial port, Ostia started to become a seaside resort for wealthy people (Aldrete 2004). During AD 62, Portus lost over 200 ships within the Claudian harbour due to a storm. This resulted in one of the most outstanding elements within Portus Port: the Trajanic basin. This basin was built to act as an harbour so the ships could navigate safely and be protected from natural phenomena (Aldrete 2004).
In the case of this project and taking into account this information, the plan of the site produced by the University of Southampton (Keay 2009) was utilised to establish essential dimensions of the building and create a 3D object depicting the building of interest. The defined size of the top plan of the building yielded adequate information to define the fundamental object. After combining the information from the three dimensional GPR interpretation vector data in which the elevation of the model is expressed (Ogden 2009) within the original plan, the basic cubic space of the object was generated using Vectorwoks, a 3D BIM software.

To subdivide the different spaces and proportions of the objects that create the building, the Fibonacci Number sequence was applied. The use of the Fibonacci Number in historical Roman and Greek buildings is present throughout history in almost every city (Elam 2001). It is through this method that every element within the model was produced in direct proportion to the essential “central square” of the building. The model contains doors with architectural ornamentation system called fornix, which consists of a simple row of arches around the frame.

6. Texture Interpretation

Under the command of Emperor Nero, Rome was enormously affected by what is now known as the Great Fire of AD 64 (Aldrete 2004). This event changed the architecture of the Roman Empire. Romans realised that the stones used to build their own houses could burn. They used bricks as an alternative and, later on, these became the standard for industrial places like Portus or Ostia (Kleiner 2009). The use of colour was more common in higher-class cities like Pompeii or Rome. Ostia is indeed known as the “Black and White City” (Kleiner 2009). Most of the mosaics of Ostia present themes related to the sea. Items like dolphins, Neptune, and ships are the most commonly used elements. It is through these mosaics that we discover evidence of lifestyles of the period.

Ostia presents murals and flooring with mosaics the same as Rome and Pompeii. The difference is that Ostia people presented all their mosaics in black and white. This is because Ostia, in common with Portus, was a working-class city (Kleiner 2009). The use of colour was more common in higher-class cities like Pompeii or Rome. Ostia presents murals and flooring with mosaics the same as Rome and Pompeii. The difference is that Ostia people presented all their mosaics in black and white. This is because Ostia, in common with Portus, was a working-class city (Kleiner 2009). The use of colour was more common in higher-class cities like Pompeii or Rome. Ostia is indeed known as the “Black and White City” (Kleiner 2009). Most of the mosaics of Ostia present themes related to the sea. Items like dolphins, Neptune, and ships are the most commonly used elements. It is through these mosaics that we discover evidence of lifestyles of the period.

Based on this information, the renders only present black and white imagery for the mosaics. This was decided presuming that Portus’ proximity to Ostia may imply a contextual similarity in both places.
7. Modelling Techniques

7.1 3D Building

As said before, the first step in the project was to create a 3D model of the warehouse. The modelling of the building needed to be adapted to the size of the plan. For this reason a complete understanding of the space was required. Combining the plans from Lanciani, Lugli & Filibeck, Testaguzza from the Portus Website of University of Southampton (Keay 2009) and Google Maps, a general plan was created in order to give a spatial context to the buildings. The model was produced depicting a multi-storey insular building with many details, expressing the specific building techniques from that specific period.

Vectorworks was the chosen software for the development of the architectural model. Although not open source, this software allows the user still to collaborate with other platforms such as Autocad with an optimised interface. Another benefit is that Vectorworks allows the user to work in 2D and 3D at the same time, accelerating the design and production process.

The model created contained essential architectural elements such as walls, floors, ceilings, doors and their frames. This allowed the model to be utilized in other applications outside Vectorworks. In the case of the 3D model, it was built with millimetre accuracy. Due to this, the 3D model contained more details than the 3D print in which greater details were omitted.

7.2 Texturing and Rendering

Once the original shapes were generated, the next process within the methodology was to texturize the objects. For this process, we looked to create visual interpretations of graphics and colour based on archaeological and historical data as mentioned in the section of ‘Texture interpretation’.

The textures needed to be mapped on the object through rendering. Rendering is the process in which geometry, lights, materials and position of the camera is synthesized image or sequence of images (Raghavachary 2005). Most of the modelling applications provide rendering solutions. On the other hand, there are specialist-rendering packages like VRay, RenderMan, FormZ, Virtus, Mental Ray or Artlantis among others. Due to the use of Vectorworks for the development of the elemental shapes for the model, Artlantis represented the fastest and most accurate solution for the rendering process, providing direct linkage and file update from modeller to renderer.

As previously said, the elemental textures are based as a set of bricks to generate the exposed brick surface. Since the exposed brick needed to present some erosion, a second texture was embedded to the walls. This was a texture of a sandstone ground applied with some basic bump to generate the texture. Instead of creating holes or a physical (modelled) transformation, the bump technique uses the intensity of the raster image to create a virtual offset on the base of the face.
of the texture (RenderPlus 2009), generating in this way a virtual physical transformation on the model. The main objective of avoiding using these types of techniques is to reduce the polygon count of the model, allowing the object to load faster into other applications. The images processed through Artlantis can be post-produced through Photoshop or any other photo retouching software in order to enhance the texture colour and contrast.

Once the 3D modelling and texturing were ready, the next step was to prepare the installation. The installation required a solid object where the textures would be projected in order to represent several interpretations or several textures. For this, once the 3D model was built there was a need to re-scale it completely. In this specific case, the model was taken down to the scale 1:00025. This was important because at a smaller scale the detail would be virtually unnoticeable. This could also generate structural problems for the generation of the 3D print as described in the next section.

7.3 3D Print

There are several 3D printing processes from which the user can choose, depending on the budget and level of detail required. This project used the ZPrinter 450 by ZCorp. This type of printer uses a powder-based material, which is soaked in a liquid based polymer. The combination of materials, transforms into a solid once they become in contact with the light produced by a high resolution Digital Light Processor projector (ZCorporation 2011). There are other types of 3D printers which use a process called stereo lithography. Experimenting with both processes, it was concluded that the costing difference and time of manufacture was more accessible using powder-based printers like the ZPrinter.

ZPrinters uses file types like STL, VRML, PLY, 3DS, ZPR (ZCorporation 2011). This project worked with STL (Alpha_Prototypes) formats since they have been generated specifically for 3D printing (Autodesk 2011). Another advantage of working with STL files is that the model can be re-scaled in resolution to reduce its file size, accelerating processing time and that is reflected in the costing of the production of the prototype (Autodesk 2011; Alpha_Prototypes 2011).

The generation of the model for prototyping had to be pre-visualised before sending it to print. It was noticed that re-scaling the architectural building to a 1:00025 also
reduced dramatically the thickness of several parts of the model. The ZPrinter 450, produces a layer thickness in between 0.089 - 0.102 mm (ZCorporation 2011). This means that many elements became virtually unprintable. To solve this issue, it was required to re-structure the model just for prototyping purposes. In this case, VectorWorks provides a protrusion and cut-out tool to add strength to the areas that may required it. At the same time, with the rescaling of the elemental structures, some of the details had to be removed due to the result of the scale of the object.

8. Installation

8.1 Technology

The installation involved several technology elements that complemented each other to provide the best interpretation. One of the technologies used was Projection Mapping, which constitute a process of projecting images to assemble them onto a substrate. In this case the substrate was the 3D print model of the building.

There has been already several cases in which projection mapping has been implemented in art projects. These projections have been used very commonly with large buildings. Examples of this are the projects developed by the company Mr Beam (Mr.Beam 2011). There have been other cases in which real size objects have been used. Examples of this is the projected face on to a mannequin head depicting a realistic face with movement displayed on the Exeter’s Underground Passages Museum, or the Wax Museum in Mexico City.

The implementation of projection mapping on to 3D printed models can allow the generation of several interpretations of the same object, helping to solve or to address specific questions in relation to time, occupation or even construction techniques. This type of technology allows the project to keep evolving without investing again in equipment, through simply updating the graphics to be presented.

In this specific case of the Portus Project, the lack of evidence makes the labour of the archaeologist more difficult. The interpretation
of the site could differ in between archaeologists or researchers. For this reason, these methods allow scholars not only to provide and share their interpretation of the same object, but also to compare interpretations.

To project the interpretation textures onto the 3D building prototype, the installation used a DLP projector. DLP projectors are less powerful than LCD but for the sake of aesthetics, DLP provides a smaller but easier to manage projector. The actual model used was the Optoma Pico a 10cm by 5cm projector mounted on a flexible GorillaPod tripod. There were other LCD and DLP projectors tested in which the Optoma presented the best overall solution. Another element that the Optoma brings to the project is its low cost value. All these projectors work with a standard VGA cable that allow any standard computer or even tablets to be connected and used for the installation.

The interface to manage the projections was developed with Flash technology. This allows easy access to change colours, interpretations or any other embedded elements. Once all these elements were integrated, the result was an installation showing the main plan of Portus, the 3D printed building and the projector that will be projecting the interpretations of the warehouse on to the 3D print. Although only one warehouse was printed, an additional important element considered was the possibility of representing the full site integrating the other buildings.

For the augmented reality there were two other models of buildings created for the installation. In this context, augmented reality has the advantage of allowing to embed an unlimited amount of models. In this manner, when archaeologists present new information, a new model can be ‘mounted’ as an alternative providing various interpretations. In this example, the project presented a large warehouse and another insular building to both sides of the 3D printed building prototype. Not only the dimension of the building can be compared in relation to the other buildings, but also the architectural standard for the archaeological site.

Another augmented reality element within the installation were two human characters used for the rendered elements. Originally this project was focused on presenting only the architectural elements. These were brought in at a later stage since integration of human figures can provide the audience a better understanding of the scale of the site. One is a female Roman character and the other one is a soldier. Both characters are presented within independent augmented reality markers. This means that the models can be presented individually, together, or even along the side of the buildings. Augmented reality allows the characters to be moved or change position since they were loaded as 3D objects, which can be seen from any angle in a screen.

There are several companies providing Augmented Reality solutions. Most of them charge for renderers, visualizers and plug-ins (Inglobe-Technlogies 2012). The objective of this project was to produce resources available to any institution or individual. It was opted to choose Flash as the main interface, where the Spark library, a full augmented reality script can be implemented to enable the camera to recognize the markers needed to mount the 3D object (Beinteractive.org 2009). In order to use real 3D objects on flash it is required to use the Papervision3D libraries (Papervision 2010). This way, the project was capable of loading a

![Figure 7. 3D object on the plane.](image)
Installation for Interpretation of Archaeological Sites. The Portus Visualisation Project
Javier Pereda

3D object through Papervision3D and mapping it on the marker with Spark. The ActionScript 3.0 libraries are open source libraries and there are other alternatives to avoid using the Adobe platform. The flash interface produces a SWF file which player is completely free to distribute and can be downloaded from the Adobe website.

9. Interactivity and Engagement

The idea behind the use of technology has been to create an installation as immersive as possible. Due to the fact that information and interpretations of archaeological sites can evolve, this kind of project has to be as receptive as possible. This is to say that there is a need to be in constant communication with the disciplines involved and to receive feedback.

For that, the installation was created with users in mind, allowing them to explore the installation at will. As said before, Augmented Reality technology works based on a camera that loads a 3D object targeted from a marker. This camera could be static or mobile presenting the content. In this case, the project enabled the user to control that camera. The control of the camera will probably be appealing to the audience inviting them to explore the space, enhancing the engagement of user and site.

The control uses a servomechanism controlled by a potentiometer, which is managed by an Arduino chip; it can also be programmed to be fully automated. Using Arduino and these types of components is also a low cost solution, enabling the project to keep adapting for further applications. The Arduino chip is completely open source and the mounting instructions and code can be found in several online forums (Ferguson 2012).

The characters provided for the augmented reality can be positioned anywhere within the plan and moved in any direction. This allowed the installation to ‘refresh’ its appearance after a certain amount of time. Alternatively, the characters can be presented individually for the sake of understanding the details of their costumes. This can be achieved through a different presentation screen or interface within the installation, or completely within a new presentation or context.

10. Troubleshooting

The process of developing installations with these types of technologies involved testing the available equipment in order to produce the best result. Most of the projections mapping examples available have been applied on to large surfaces like buildings or rooms. The technology available has been designed to work at a large scale and it is similar to working with big format printing. Similar to a poster that is
required to be seen from a long distance and requires a large format, its printing resolution is taken down to reduce costs and accelerate the process. The prototype has 30 cm width of projection space. This means that the scale of the projection area is smaller than intended as a regular use. This makes the user more sensible to the detail of the projection. Along this, the DLP projectors generate blurriness on the edges of the object that on a large scale avoids having big pixels on the projection. In this case creates an effect of blur on the image projected. This generated a problem with one of the interior areas. The thermopolyum projected on the left side of the building did not contain enough resolution to actually notice the architecture. Different projectors with higher resolution provided a better result but still not good enough to keep the element there. Because of this it is important to analyse and know the capabilities of the projection technologies used and define or adapt the interior elements to be able to provide a good size projection and fulfil the installation’s purpose.

In this case, increasing the scale of the prototype did not seem like viable, due to the scale of the entire plan, but a more detailed presentation of that specific building as an independent installation can provide the exploration of other aspects such more specific architectural detail.

On the other hand, modelling the objects presented many issues when loading them on to the augmented reality application. Papervision requires a Collada object, therefore it is important to provide a .dae model. These models were generated from several 3D cad applications. These applications use a fine triangulation process to generate the objects. This creates the application to process more shapes and slows the computer down resulting in frame drops. The solution in this case was to reduce the polygon count into a coarser triangulation. This will produce a less defined model or prototype that provides a faster processing result.

11. Further Development

Several applications allow augmented reality or 3D models to be rendered in real time. The project used Flash as the main technology since it is already distributed through the web and most of the computers have the Flash player installed on their browsers.

Implementing content with Unity and WebGL would probably be the next to distribute in alternative ways 3D content. Working with 3D and augmented reality for mobile has become more popular and this may expand the user experience opening opportunities for institution and individuals to implement new ideas.

Further opportunities for the project involve automating of the augmented reality cameras. Another implementation for the installation may be to use an Arduino robot with a camera mounted on the top following a path drawn on the plan of the installation showing at the same time the augmented reality markers.

Regarding modelling, most of the models were loaded as polygon-based models. It is imperative to be able to produce models that are capable of loading with ease on mobile phones to be able to create a sustainable technology capable of being applied anywhere.

12. Conclusions

The aim of this project was to use design elements as mediums of communication. Exchanging information between different disciplines (Graphic Design, Museum Studies and Archaeology) this project has been able to consolidate quite successfully. The interdisciplinary approach taken from Graphic Design allows projects like this to be implemented. The project combined disciplines including archaeology, architecture, multimedia, visual communication and education among others, attempting to
produce a tool that helps visualizing diverse representations through an intuitive installation that can be implemented anywhere.

The line where the duty of each specialty finishes has become less visible as technology and industries evolve. It is difficult to pre-plan when to bring specialists from other areas into the field that is being developed will be required. The familiarization of specialists with these types of projects can promote other disciplines to become more open to collaboration. However, one of the main obstacles implementing technology is usually monetary. For this reason the project has pursued the idea of utilizing as much as ‘open source’ material as possible to motivate researchers to experiment with these kinds of technology and promoting interdisciplinary research.

References


Material Motion: Motion Analysis for Virtual Heritage Reconstruction

Kirk Woolford
University of Sussex, UK

Stuart Dunn
Kings College London, UK

Abstract:
Through the AHRC funded “Motion in Place Platform” project, a number of experiments were conducted to look for quantitative differences in movement in virtual vs material environments. Actors were asked to enact a number of activities hypothesised to have occurred in a British Iron Age roundhouse while wearing inertial motion capture suits. These activities were recorded both in a “virtual” studio (re)construction as well as material (re)construction at Butser Ancient Farm. The data from these experiments was then analysed to look for differences in movement which could be attributed to artefacts and/or environments. This paper explains the structure of the experiments, how the data was generated, how it has been analysed, and what theories may make sense of the data and what conclusions have been drawn about how objects and environments may influence human movement and how a better understanding of movement many help understand empirical remains.

Keywords:
Motion Capture Data, Reconstruction, Virtual Reality, Experimental Archaeology

1. Introduction

In British Iron Age domestic culture, there are no historic or material referents to how particular houses were built, how they were used, or how artefacts such as arrowheads or ceramics were made. The corresponding methods must be inferred by a process of logical deduction, and examination of the available empirical evidence. However, how we approach this process of deduction can, and often does, involve a human factor. The reconstruction process in experimental archaeology now has a long tradition of researching and utilizing past methods construction and craft to construct (the term ‘reconstruct’ is explicitly avoided in the literature – see Reynolds 1993) non-extant buildings using those methods. The experimental approach, now well established and widely referred to, requires the ‘human factor’, in that it requires human intervention in, and interaction with, the physical world. While it is not possible to go back in time to capture the exact motions involved in archaeologically relevant activities, we can capture current activities and the physical processes in order to gain more insight into probable past activities.

Experimental archaeology is a branch of archaeology, which replicates or attempts to replicate past processes in order to understand what is found in archaeological record. This branch is often cited as offering an important asset in the study of human interaction with material culture, especially when dealing with remote periods of history where there are few othersourcesofdataonthehumaninterventions. However, due to an understandable desire to adhere to empirical evidence, means of inferring the human movement behind interventions are rarely considered in the reconstruction of archaeological environments. The most obvious reason for this is that buildings, features and artefacts can be understood and reconstructed (whether digitally or not) from empirical archaeological remains, whereas there is little or no direct evidence for how
people might have looked and moved through the places they created. Approaches that seek to go beyond this are methodologically fraught as a result of ‘the human factor’. It is further inevitable that such living interpretation will be problematic, since environments, objects and landscapes are, to one extent or another, cultural constructs: society attaches significance to landmarks and features which cannot be retrieved without written records. However, implicit in all archaeological interpretation is the truth that this human factor is behind the process of the material record’s creation. Human processes have, in the past, been regarded as intangible and unrecoverable, and therefore implicitly and explicitly written off in experimental archaeology. For this reason, experimental archaeologists have traditionally shunned ‘the human factor’, focusing instead on the re-creation of archaeological features from empirical evidence (Harding 2009; Coles 1979). Indeed, the very notion of attempting to include ‘the human factor’ in experimental reconstructions is viewed with scepticism at best and outright hostility at worst. As Peter Reynolds, the founding director of the Butser Ancient Farm project has put it:

“In real terms it is only sensible to examine structures physically and as far as possible to dehumanise the examination process. Re-enactment is best left as a dramatic indulgence to the imagination, which can be recognised as singularly valueless and instantly forgettable ... History, and by implication prehistory, is swiftly becoming a tabloid newspaper sub-editor’s view of the past” (Reynolds 1993).

While some might view the strength of this distinction as being somewhat harsh, it nonetheless highlights a significant gap not only between ‘physical world’ reconstruction projects such as Butser Ancient Farm and cognitive considerations of how the reconstructed spaces may have been used, but also between the application of virtual reality reconstruction and any attempt to (re)create past movements in any place, physical or virtual.

The Motion in Place Platform (MiPP) was developed through cross-disciplinary collaboration between researchers from the Universities of Sussex, Bedfordshire, Chichester, Reading, Kings College London, and others in order to study relationships between human movement and place. The platform itself is described elsewhere (Dunn et al. 2012) and a great deal of information is to be found on the project website (http://www.motioninplace.org). This paper focuses on applications of the platform to the understanding of human movement in virtual heritage reconstructions. Specifically, the paper examines relationships between movement and artefact, movement and space (virtual or physical) and differences between expert/amateur or informed/uninformed movement. The paper presents some of the quantative data produced through a series of experiments with differing artefacts, spaces, and actors.

2. The Movement of Things: Relationships Between Movement and Artefact

The Motion in Place Platform (MiPP) was developed through cross-disciplinary collaboration between researchers from the Universities of Sussex, Bedfordshire, Chichester, Reading, Kings College London, and others in order to study relationships between human movement and place. The platform itself is described elsewhere (Dunn et al. 2012) and a great deal of information is to be found on the project website (http://www.motioninplace.org). This paper focuses on applications of the platform to the understanding of human movement in virtual heritage reconstructions. Specifically, the paper examines relationships between movement and artefact, movement and space (virtual or physical) and differences between expert/amateur or informed/uninformed movement. The paper presents some of the quantative data produced through a series of experiments with differing artefacts, spaces, and actors.

2. The Movement of Things: Relationships Between Movement and Artefact

The opportunities for testing and experimenting with motion capture technology across both ‘real world’ and VR settings are rather limited, given the relative rarity of 1:1 scale 3D reconstructions of non-extant historic sites. The Butser Farm is an example of such a reconstruction. However, an experiment in which motion data from trials at Butser which could be compared with those from a virtual environment obviously required an equivalent virtual reconstruction from independent data. During their Summer 2010 excavation of the Silchester Roman Town [6], evidence was emerging of an earlier Iron Age town on the Silchester site. One feature of this evidence was a clear circular contour of a structure. As more evidence arose of an Iron Age town, interpretations about this evidence included this structure, from which was inferred the existence of a round house on the site. The MiPP team agreed to build a virtual (re)construction
of an Iron Age roundhouse according to the dimensions found at Silchester, and to populate it with virtual characters conducting activities which are likely to have occurred in such a place and time, and which reflect the likely domestic character of Iron Age round houses. In order to capture movement for these virtual characters, the team followed the norms used for film and video game motion capture. The location and boundaries and of conjectured walls, hearth, and other elements were taped onto an empty stage floor. Various items were used as stand-ins for artefacts, and dancers were given direction by a choreographer or movement coach as to how they should move and what they should do while wearing the motion capture suits. The dancers were asked to perform various ‘everyday’ tasks such as fetching water and wood, tending a fire, and sweeping the house. The value and validity of these movements was not questioned until one of the dancers picked up a modern push-broom from the corner of the studio and began sweeping with it. It was immediately obvious that this was a very specific movement closely linked to an artefact which would not have been present in an Iron Age roundhouse. After the dancer was reminded that the push broom was a 20th century invention, she swung the broom from side-to-side without touching the floor. Neither of these sweeping actions could be considered “correct”, neither helped understand how or why round houses were constructed in forms we have found, nor did they illustrate how these structures and artefacts were used.

This event illustrated the difficulty of using contemporary motion capture techniques developed for film and video games for virtual heritage modelling. Most motion capture tools used in virtual heritage were developed for the entertainment industry where the “look” and “flow” of the movement is more important than its provenance or ‘accuracy’. Many motion capture tools exist for medical or biomechanics applications, but these are seldom used. More importantly, the experience documented the impact the artefact (i.e. broom) had on the movement. This is illustrated in figure 1 examining the physical stance and by tracing the path of the user’s left hand while sweeping in a studio with both a contemporary push broom and a broom approximating ones likely to have been used in the Iron Age.

This link between specific artefacts and the movement required to manipulate them is well known. Its importance is clearly explained by French sociologist Marcel Mauss as he described cultural difficulties with techniques required to shovel soil:

“during the War I was able to make many observations on this specificity of techniques. e.g. the technique of digging. The English troops I was with did not know how to use French spades, which forced us to change 8,000 spades a division when we relieved a French division, and vice versa. This plainly shows that a manual knack can only be learnt slowly. Every technique properly so-called has its own form.” (Mauss 1973, 71)

The Norwegian archaeologist, Bjørnar Olsen points out the necessity of understanding the material culture of “things”:

“Why is it that things, the material world, have escaped the attention of the
contemporary social and human sciences? One reason frequently given is that things do not call attention to themselves — they are so integrated in our lives, being at the same time the ‘most obvious and the best hidden’” (Olson, 2003, 94).

In all of our daily conduct, objects are involved as (more or less) taken-for-granted and inherent aspects of our doings. They do not just provide frames, scenes, or background for our actions, but are intrinsically and indispensably involved in enabling those very actions. Thus, the time seems overdue to credit them some social recognition. This is not to say, of course, that things are the only vital component of social order and constitute the site where all our attention from now on should be focused...

Guided by the long-held concern with signaling and identity formation (status, gender, ethnicity, personhood, etc.), the material cultural focus has been primarily related directly to bodily display and inscriptions (ornaments, dress, tattooing) and iconic manifestations such as figurines, masks, anthropomorphic rock art, and so on. Despite the claim that “under the influence of phenomenological approaches” the focus has shifted to analyses of “the production and experience of lived bodies” ... the pivotal role of the human being is rarely challenged. Rather than exploring the possibilities opened by focusing on somatic experiencing and, consequently, on being as a materially entangled being, many archaeologies of the body may actually be seen as reinforcing the anthropocentric bias... This adopted bias accentuates my initial claim that archaeological theorizing should make a difference by always and consistently remembering things.” (Olson 2010, 18)

Such understandings are (further) complicated in terms of their interpretation in the present, when the movements of interest occurred in the past – especially the distant past. The philosophical construct of phenomenology, of interpreting locations in terms of experience of them, has a long heritage in archaeology. Typically it has focused on the embodiment of interpretation of locations-specific practices such as cult and religion, or the remediation of pathways through the landscape that are demarked by some extant physical structure, such as earthworks (see Tilley 1994; Copeland 2009). In a wide-ranging review of the subject in 2005, Joanna Brück notes that"

“[O]ne of the most productive strands of phenomenological writing within archaeology has been the deconstruction of the dualistic thinking that is a product of post-Enlightenment rationalism. This has facilitated a radical reconceptualization of the nature of materiality and the relationship between people and artefacts. ... Only by seeing objects as inanimate can we adhere to a model according to which humans impose meaning on a passive and pre-cultural universe. If, on the other hand, we recognize that artefacts, buildings, monuments and landscapes not only affect us but make us who we are, then our engagement with the archaeological record is necessarily a dialogue in which both archaeologists and the axes, houses or burials we study are created and transformed” (Brück 2005, 65).

3. Motions in Place

The experience with the broom showed that the connection to material objects such as tools and buildings are of crucial importance in elucidating our understanding of possible behaviours and movements at a historically inaccessible period. Consequently, a further set of experiments was devised in an attempt to test the influence of a place or location on movement. These experiments focused on the task of sweeping within two (re)constructions of the same round house. Both round houses were constructed according to excavation data from Moel y Gerddi, Wales (see http://www.buterancientfarm.co.uk/pdf/moel%20y%20gerddi%20guide.pdf). The first, immaterial, or virtual round house was created using
projections and a head-mounted display at the University of Sussex. The second, material, or physical round house was constructed of materials expected to have been available in Iron Age Wales at the Butser Ancient Farm in Hampshire, England (Fig. 2 and www.butserancientfarm.co.uk).

Two dancer/choreographers were given a broom, constructed using materials and methods sufficiently generic as to approximate to those likely to have been used in the Iron Age, to sweep the virtual round house as well as the physical round house. (Fig. 3). In the virtual round house, their movements had no effect on the virtual environment. The smooth, flat floor of the studio offered little resistance to the brooms and the even floor and lack of physical consequences related to sweeping through posts or walking into walls appeared to invite the dancers to move aggressively and openly. In the physical round house, the floor was uneven and the dancers had to move the broom around posts while not stepping into the hearth. There was great deal of variation in the resistance to the movement of the broom on the floor. At the same time, the dancers learned that large, fast movements created dense clouds of dust and damaged the floor of the house.

In order to analyse the capture data created in both versions of the roundhouse, the authors developed a bespoke application to track the position of the dancer’s hands while sweeping and to determine the distance the hands travelled and the amount of time required for an average “sweeping” motion or cycle. A single sweep motion or cycle was defined as the time between when a broom was placed down on the floor until the next time it was placed on the floor. Figure 4 shows a plot of sweeping in both the virtual roundhouse (4a) and the physical roundhouse (4b). Both graphs show the position of the dancer’s right hand over approximately 45 seconds of sweeping. The plots in the bottom right show the composite 3D motion trajectories the hand (i.e., its position in 3D space). The other two graphs plot the distance away from the centre of the body. The top graphs show these positions on a traditional timeline while the graph in the bottom left plots y-offset, (the height above the body’s centre) on the y-axis against xy-offset (the length of a vector from the center of the body to the body part being tracked). This plot also highlights the current sweep cycle or stroke and the current position in this cycle.

The numbers listed in the bottom-left graph indicate the duration of the current sweep stroke, the distance travelled by the hand, and a numerical representation of the smoothness of the stroke. By averaging the durations and distances of all sweep strokes, the following were determined:

4a., virtually (re)constructed round house:
- avg stroke dur, 2.3 sec,
- avg stroke dist 8.97 cm/sec.
What does this mean? This data would appear to demonstrate that the dancer did, indeed, make larger sweeping strokes in the virtual roundhouse (as expected). However, the dancer also made sweeping strokes of shorter duration in the physically reconstructed roundhouse. This may be a result of the dust stirred up by sweeping in the physically constructed space, or it may be a result of the amount of resistance of the rough, uneven floor. Because the sample size is so small, it’s not possible to make any definitive statements, but the data does appear to demonstrate that engagement with the environment has altered the dancer’s movement.

This coincides with much writing on movement and environments as summed up by the Architecture theorist, Juhani Pallasmaa:

“Our bodies and movements are in constant interaction with the environment; the world and the self inform and redefine each other constantly. The percept of the body and the image of the world turn into one single continuous existential experience; there is no body separate from its domicile in space, and there is no space unrelated to the unconscious image of the perceiving self.” (Pallasmaa, 2009, 40-41).

4. Informed Motion

The motion experiments detailed in the past 2 sections were conducted with dancers. Dancers were used not because of their virtuosic movement abilities or vocabularies, but because of their ability to take physical direction, remember and re-create the movements. However, when working on site at Butser ancient farm, the MiPP team were able to capture the movement of a number of the experimental archaeologists working on the site. In addition, the dancers were captured upon first arriving on site, then captured again after having been given training by the archaeologists who worked on the site on a daily basis, performing the same tasks. The dancers’ movements were then compared against the archaeologist’s movement and their earlier, uninformed motion as depicted in figures 5a, 5b and 5c.

This is, in effect, an extension of experimental archaeology, which allows us to infer how people are likely to have interacted with their physical environments and how those environments (or tools) were constructed. It also resonates with Marcel Mauss’ theory of techniques of the body, transmitted through tradition:

“I call technique an action which is effective and traditional … There is no technique and no transmission in the absence of tradition. This above all is what distinguishes man from the animals: the transmission of his techniques … we are dealing with techniques of the body. The body is man’s first and most natural instrument. Or more accurately, not to speak of instruments, man’s first and most natural technical object, and at the same time technical means, is his body.” (Mauss 1973, 73)
5. Conclusions

As noted, the purpose of this exercise is, emphatically, not to attempt to re-enact possible scenarios of history or prehistory, but to capture and visualize human interaction with place and material culture as documented by archaeological evidence; and thus to provide a critique of how well VR represents the experiential past. No, it is not possible to definitively know how Iron Age Britons used their roundhouses. We can infer past movements from an understanding and analysis of current movement in much the same way we infer the structure of past buildings and material objects through the fragments that have survived to our current time. However, just as archaeologists make clear distinctions between what material objects have actually been uncovered and what contextual information they have based their conjectures upon, we need to be clear about exactly what motion data we are capturing and the contexts in which it has been captured. If we want to understand how motion influences place and place influences motion, we need to capture and study them together.

Acknowledgements

The platform has been developed by a cross-disciplinary team including Stuart Dunn (Centre for e-Research, King’s College London), Mark Hedges (Centre for e-Research, King’s College London), Helen Bailey (Centre for Applied Research in Dance, Bedfordshire), Sally Jane Norman (Attenborough Centre for the Creative Arts, Sussex), Martin White (Computer Graphics Centre, Sussex), Sarah Rubidge (Dance, Chichester) and directed by Kirk Woolford (Media, Film and Music, University of Sussex) and supported by Leon Barker and Milo Taylor. MiPP is funded by the U.K. Arts and Humanities Research Council under the Digital Equipment and Databases for Impact (DEDEFI) Scheme. For more information, please see http://www.motioninplace.org.

References


Copeland, T. 2004. “Presenting archaeology to the public:


Interactive Workspace for Exploring Heterogeneous Data

Uros Damnjanovic and Sorin Hermon
The Cyprus Institute, Cyprus

Abstract:
In the vast sea of available digital data, it is getting more and more difficult to access the right piece of information and make some practical use out of it. Information overload is a term that describes the situation when users are confronted with huge amounts of content and are getting overwhelmed with it. This paper presents a part of our work on developing new means of interacting with data in digital environments. The main goal of our work is to explore the possibilities for introducing new interaction tasks and new interaction models in digital information repositories. Our framework enables users to manipulate and adjust visual layout of the searching task, and to easily compare results of different searches. We also show how individual searches may be stored and used in exploratory data analysis by comparing elements of individual searches and corresponding data.

Keywords:
Digital Libraries, Human Computer Interaction, User Interfaces, Information Retrieval, Content Management

1. Introduction

We are all aware that in today’s digital age information is being created and used at a staggering pace. Availability of numerous devices for creating and sharing information drastically changed channels of communication over the last two decades. In the vast sea of digital data, it is getting more and more difficult to access the right piece of information and make some practical use out of it. Information overload is a term that describes the situation where users are confronted with huge amounts of information and not being able to make sense out of it. In order to help users access, organize and use information, there is an ever-growing need for new interaction tools. These tools should help users identify and access useful information and transform it into knowledge. This paper presents part of our work on developing new means of interacting with data in digital repositories of archaeological data. Whenever there is a digital content that hold some useful information, there is also a need to explore and interact with this content in order to make practical use of available information. Conventional systems for accessing digital information provide only basic possibilities for search and exploration of the content. Accessing content by means of simple key-word search and document listing, as it is done in most of the currently available digital repositories, is definitely not enough to make practical use of data. Users need to be able to explore and compare data, be able to explore relations between individual items, and to access data in a way that suit their needs. User interfaces should help them in these tasks instead of being a constraint.

The main goal of our work is to explore the possibilities for introducing new interaction tasks and models in digital heritage repositories that will help users explore and use available information. By interacting with the content in different ways, users have a chance to understand complex relations which are not obvious when simply browsing the repository or when using a key word search. More ways of interaction with data could increase the probability of new knowledge being created from the available information. Each interaction task can be seen as a different angle from which data is being observed. Modern digital repositories should stimulate understanding,
sense making, knowledge creation and critical thinking, by providing various interactive tools for exploring and accessing data. We argue that intuitive interaction with the content can improve communication and knowledge sharing potential of information systems. Innovations in Human Computer Interaction research, particularly novel interaction techniques, are still rarely incorporated into products (Beaudouin-Lafon 2004). It is argued in (Jordan 2010) that the most important and at the same time the most difficult task in digital information systems is to make the step from perception of the information need to a request that can be used to query an information retrieval system. Available information retrieval systems restrict the degree of freedom of the user in this aspect: often they only support a few search activities well and other search activities have to be performed using external means for support, making the search less integrated and less pleasant for the user. When users have poorly defined or complex goals, search interfaces that offer only keyword-searching facilities provide inadequate support to help them reach their information-seeking objectives. The emergence of interfaces with more advanced capabilities, such as faceted browsing and result clustering, can go some way toward addressing such problems (Wilson et al. 2009)

We present a framework for interacting with information rich digital systems such as digital libraries, digital museum collections or e-learning systems. We focused our work on developing interactive users interfaces that provide users with high level of freedom when interacting with the content, and also provide users with a large number of possibilities on how to explore data. We wanted to find solutions for practical problems such as formulating visual search queries, presenting and exploring the set of results, comparing search results, selecting items for further investigation, managing large number of searches and search results. All these tasks should help users in various stages of interacting with content. However our focus was on the content creation stage, where the goal is to create data collections that will cover specific topics, and on exploring and understanding knowledge embedded in these collections.

Commonly, a search task is understood as a tool for helping users find some specific piece of information. In our approach it is a tool for helping users in their exploration of the content, by formulating visual queries and exploring relations between queries and data. Studies show that about 50% of Web search is for information exploration purposes, where the user would like to investigate, compare, evaluate, and synthesize multiple relevant results. Due to the absence of general tools that can effectively analyze and differentiate multiple results, a user has to manually read and comprehend potential large results in an exploratory search (Liu and Chen 2012).

Our framework is focusing on cases, when users perform searches not only to find specific piece of information, but also to explore and investigate the structure of data, and distribution of various properties in a dataset. In order to make this possible, the process of query formulation and search result exploration needed to be redesigned, and a new interaction tasks provided that will help users achieve their goals. Our interactive search framework has a goal to help users explore data by formulating all kinds of queries and by being able to interact with various aspects of data, by comparing different search results, by comparing quires and relevant data sets and by comparing individual items.

2. Interactive Framework for Accessing Digital Information

There are many different motives for accessing and using digital information. Even though there are different motives and potential scenarios for use of digital information, there are few common things for all these tasks.
The essence of these tasks is to access and consume digital information in order to satisfy information needs. Our framework is intended to help users explore digital content, and understand available data by exploring its properties. Whenever there is a need to generate digital content in a way that will enable various applications over generated data, there is also an obvious need for a content management tool. A content management tool gives content creators full control over the published data. This means that the information can be easily removed or added to the system, modified and linked to each other. Digital collections of various data are used these days to explore various aspects of our life, ranging from art and history to science, fun or business. Archaeology is not an exemption on this. In our research group we initiated efforts on building a digital repository of archaeological data that will provide access to data both to the members of the group as well as to the general public. As a part of everyday work on various projects, large amount of digital content is created on an everyday basis. These include various types of documents such as images, texts, transcripts, 3D models, videos and so on. As the amount of information was growing we realized the need for a single tool that will allow us not only to store information in a repository but also to explore available data. Most of the digital content created in our group is accompanied with rich metadata descriptions that provide potential for exploring data in many different ways. For this reason we decided to work on a new content management framework, that will be completely focused on enabling various kinds of interactions with data. When dealing with large amount of information one can identify two main goals. First, there is an obvious need to store data together with its descriptions in a way that will enable easy access to the content. This is the focus of most of the repositories available today. The second goal that attracts more and more attention over the last years is to learn how can these data be practically used. Instead of publishing data just for the sake of publishing, we think that it is important to provide means for practical use of this data. This calls for inclusion of end users in the development process. By interacting with end users at the early stages of a project, developers may get useful ideas on what tools should be included in a system, how users are planning to use provided information, and what should be the focus of the interaction with content.

The main goal of our work was to provide means for controlling the generation of content in our group’s repository, and to be able to learn and explore various features of our data by using richness of metadata descriptions as a basis for interaction. Our data is organized in different collections, described with different metadata schemas. When there is a need to explore distribution of various properties and its relations our framework provides an intuitive interactive interface for doing this. Another possible application of our framework is for controlling the process of content creation. When building a digital collection there is always a goal that needs to be achieved by presenting content to the users. In order to be able to convey the message correctly, digital information needs to be carefully selected and presented. Complex ideas need to be well represented and described from various aspects.

3. Understanding Data by Interacting

Interaction with digital content can be seen as a channel for accessing and understanding complex information stored in digital environments. In order to get the comprehensive view and understanding we need to be able to compare different pieces of information, to understand relations between different items, to understand importance of specific items compared to the whole group or to the other items. Sometimes it is not the data itself that is of importance, but a data property might be of interest, and understanding of how this property is distributed over the whole dataset can help users in achieving their
task. Deep understanding of relations and distribution of properties contained in a data set can be of great interest when building digital collections, or when compiling a digital learning material. When it is important that all topics of interest are covered users might wish to explore data properties rather than data itself. One example of such a tool that enables exploration of data is the Relation Browser (Capra and Marchionin 2008). It is a tool for understanding relationships between items in a collection and for exploring an information space. Users can explore datasets using a mixture of searching and browsing, supported by keyword search and dynamic queries using facets. Combination of query and browsing strategies can assist users in formulating queries, forming context for particular search and exploring and gaining comprehensive view of the collection. In another similar example (Shiri et al. 2011), authors developed two interfaces that combined search and browsing, supported dynamic exploration of conceptual structures of a thesaurus, and provided dynamic term relation features to give overview of data. Authors showed that browsing through a thesaurus improved users understanding of the relationships between materials and the catalogue resources. We look at the interaction and user interfaces as channels for information transfer between two large pools of information, digital information world and the users knowledge and experience. The goal of the advanced interaction tools would be to increase the capacity of these channels to enable a flow of a sufficient amount of information thus helping users increase their productivity in the every day tasks. Another important aspect of the interaction is that it should help users process the information and stimulate cognitive processes in users that will result in the generation of a new knowledge (Natrha et al. 2011).

In Wang et al. (2011) the authors presented an intuitive interactive interface for browsing of a large-scale image and video collections. Their approach is based on visualization of the underlying structures of a dataset and on the visualization of the size and spatial relations of displayed images. In order to achieve this, images or video key-frames are initially clustered using an unsupervised graph-based clustering algorithm. By selecting images that are hierarchically laid out on the screen, user can intuitively navigate through the collection or search for specific content. The work presented in this paper makes a shift towards a more user-centered design of the interactive image and video search and browsing interfaces by augmenting user’s interaction with content rather than learning the way users create related semantics. This shift enables not only efficient retrieval of the desired content, but offers more intuitive access to vast visual data and often gives unexpected perspective of the explored dataset. Presenting and browsing of the image search results plays a key role in helping users finding desired images from the search results. Most of the existing commercial image search engines present images based on their position in a ranked list. However, such a scheme suffers from at least two drawbacks: inconvenience for consumers to get an overview of the whole result set, and the high computation cost to find desired images from the list. In this paper, we introduce a novel search result summarization approach and exploit this approach to further propose an interactive browsing scheme (Fasr and Sedig 2011).

Interaction with data can be seen as a twofold process. One side of the process is the user’s input, and the other side is the output of the action. Mental processes in the users cognitive system are connecting these two sides, of interactions eventually leading to the generation of new ideas and of new knowledge. Smart interaction systems should make use of this, so that every interaction task have the clear goal on what should be the cognitive output in terms of understanding and processing of information. The purpose of the interaction is to help create knowledge, develop understanding and acquire insight from the resources in collection (Chen et al. 2009). Set of possible interactions is not infinite,
but is large enough to provide many different advanced possibilities for understanding and learning new information from existing content. Interaction can help transform data to information and information to knowledge (Gwizdka 2010).

4. Interactive Search

The main idea of our approach was to create an interactive space in which users can explore digital collections by being able to freely manipulate a content of a page while searching, browsing and accessing information. In order to allow flexibility without assuming preferred search style of every user, we decided to develop a search interface that will let users easily modify visual layout of a page to suit their own needs. Instead of having the predefined, fixed search interface, in our approach users build their own search interface, starting from an empty workspace by adding various search options from the search control menu. The search control menu contain all possible options for setting up and customizing the workspace, and is always visible on top of the workspace, enabling users to add or remove items easily (Fig. 1).

The options of the control menu are divided into three groups. The first group is search options, which is used for setting up search controls, and defining queries. The second group of options is control related to the result display. And finally the third group of options is related to the visual options of the workspace, helping users customize the look of their workspace. When an item is selected and added to the workspace, users can drag it and position it around the workspace. Also users can resize any item, giving them total control over the visual layout of the workspace

4.1 Visual Search Creation

In order to perform a search, a search form needs to be created from available search elements. The search form is created by selecting a type of the search from one of available options. At the present moment users can chose from simple key work search, advanced search, thesaurus search and search by example. After the type of the search is selected users can add specific options for each type of the search (Fig. 1). For example in the thesauri search they can browse a thesaurus and select terms of interest. For the advanced search, users may define, in addition to the key word, specific properties to be searched, or define constraints on any available property. In order to do so, users may select a data description from the list of available metadata formats, and add it to the search form. This means that the search algorithm will look for the selected property and will return only items that satisfy defined criteria (Fig. 1).

This is useful when dataset contains more than one data collection, possibly described with different metadata schema. In order to be able to search over all available collections, we provided a possibility to define search filters from any of the available schemas, with results being formed only from items that satisfy defined conditions independently of its metadata schema. By selecting metadata schema from the search menu, each metadata field in the schema depending on its type may be used as a filter in the search. This is done either by defining a key word query for a specific field, or by selecting
a range of values for numeric properties, or by selecting a set of values for available options. In order to be able to add various filters in the search, there needs to be a set of clearly defined metadata properties for every metadata field used to describe data. When a new data collection is added to the system, a metadata structure for this collection is defined and each property described by its type, possible range of values, or by a set of discreet textual or numerical values. Our idea was to create queries by using only search elements that are going to be used in the search. By stacking together criteria’s that are actually used in the search, instead of having all available options present in the search form, users get visual sense on their search conditions. In this way, corruption of a visual space used for exploring the content is prevented, leaving only necessary information on the screen, and helping users understand the structure of their search better. We wanted to use visual representation of the query as a simple indicator on the complexity of the search.

### 4.2 Search by Example

Another important type of a search available in our framework is a search by example. A search by example is well known functionality in content-based information retrieval, where query is analyzed, set of feature vectors calculated and used to find similar items. It is usually used in content-based image search where set of visual features is used as a search input. In our case we use a bit different approach, instead of using content-based data representation we used metadata properties to find similar items. Since every digital object will be assigned metadata description, this information can be used to search for items that share same or similar properties. In our case users may select one or more items, drag them to the search form, and then define a subset of properties that will be used in the search (Fig. 2). This will add a constraint to the search on which metadata fields are evaluated and compared to the query. If no filters are defined the search is performed on all metadata fields that are available in the query, and results will be arranged according to the level of matching. Items that have more properties matching the query will be ranked higher, while items with fewer properties matching the query will be ranked lower.

### 5. Accessing and Exploring Search Results

#### 5.1 Query Visualization

After each query has been submitted, its visual representation is automatically created for use at later stages of interaction, when users might want to explore submitted queries. Exploring search queries can be useful for many reasons: to evaluate search history, to make sure data is covered correctly in searches or to evaluate data importance levels. Queries can be explored in a sequence as they were submitted, can be grouped based on their structure and explored on a group level, or just browsed using a visual query representation. In every example a visual representation of a query is used to represent the query. Visual representations are automatically created after each query is submitted, by assigning
specific symbol for each search type. Together with search type symbol, depending on the search conditions, each query representation is assigned additional visual representation showing how many properties are used in the search and type of the properties. Visual representation doesn’t show all information about the query rather it shows the structure of the query and the type of the query. I needed users can always get full information about the query. Another possibility for managing and exploring submitted queries is that users can submit new search not over the data, but over the set of submitted queries. In this way users can search and explore their queries, and use them to get better view of their activities, and as an input for new searches.

When running multiple searches, many times users want to be able to compare results of different queries, and be able to modify search criteria’s accordingly. Using the conventional search engines this is not an easy task to do, since every search will lead to a page with results, and in order to run another search and be able to compare results, one needs to access another search window, run the search and then compare results in two different windows. This is neither practical nor efficient, since users need to spend their time on arranging windows on the screen, comparing items from different windows and use acquired information on another window. All this may result in a cognitive overload, state in which users ability to understand presented information is significantly reduced due to the users engagement with non important activities such as arranging windows, switching from one window to another, memorizing different queries. In order to overcome these issues we used different approach to the way a search is done. In our approach users may run as many searches as they like in the single screen. Results of each search can be displayed either in the single result panel, or in separate panels depending on users preferences. If users select the option that each search opens up a new result panel, users may easily arrange position of result panels in the screen. They can move panels around the screen and adjust their appearance. Users can also modify how images are displayed in each of these panels by zooming in and out, adjusting the size of image thumbnails and levels of details about each image. By clicking on an image in any of the resulting panels new panel with image details is opened.

One of the main goals of our framework was to help users achieve their tasks by improving the overall interactivity of the framework. It means that every image from resulting panel can be used as an input for a search by example. Example query is defined by dragging one or more images to a designated area of the search form. After defining query objects, the user can define which fields of selected objects will be used in the search. Another interaction task that can be done in the workspace is to use metadata descriptions of each individual object to form a new query. When exploring object details in the separate window, users might find features of interest, which can be selected and then added to a new query. This saves time and effort of memorizing, selecting and typing various properties and enable users to efficiently navigate through the information space. Our goal was to make all the information representations as much adjustable as possible. In the same way as an image is dragged and used in the query by example, each metadata property can be used in the search. As we have already mentioned one of the options in our framework was to explore and organize past queries. Another way of exploring past queries is by selecting its result panel. When every search returns results in a separate search panel, we needed a way to be able to easily associate each result panel with its query. Since there is only one query panel and there may be many result panels, by clicking on a result panel query panel is automatically updated to show corresponding query. This provides easy way of matching query to the selected data set, and opens up possibilities for comparing queries and respective results (Figs 3-4).
to add new bookmark panel in the workspace, and then add items from any search into the bookmark panel. It adds another data panel to the workspace that can be explored and accessed as any other search result panels. In addition, by clicking on an image in the bookmark panel, search query panel is updated to represent the query that is used to bookmark selected image. By selecting an image from the result panel, query that resulted in selected dataset is stored in the system to enable users to explore queries related to specific images (Fig. 5).

6. Conclusions

In this paper we presented our efforts on developing an interactive space for controlling and exploring content of digital repositories. Our work is focused on creating new interaction tasks in digital repositories that will give users chance to explore the content in new and efficient ways. Presented work is a part of our effort to build digital library of archaeological data coming from various projects in which our research group participated. In order to be able to manage large quantities of information coming in various formats and from various sources, we needed solutions that will not only provide access to information but will help users explore and understand available information. Initial ideas for the project were result of collaboration between computer scientists, which were in charge of developing and maintaining of the repository, and archaeologists that were generating and using the content from the repository. From this collaboration set of potentially interesting tasks emerged which were used as a basis for our interactive content management system.

References


Simulating the Past
The Use of CFD to Understand Thermal Environments Inside Roman Baths: A Transdisciplinary Approach

Taylor Oetelaar  
University of Calgary, Canada  

Clifton Johnston  
Dalhousie University, Canada  

David Wood, Lisa Hughes and John Humphrey  
University of Calgary, Canada  

Abstract:  
The engineering technique of computational fluid dynamics (CFD) can provide another tool for the classical archaeologist to use when analyzing ancient structures. CFD can create an approximation of the temperature distribution and the velocity profiles based on the geometry, materials, and environmental conditions that are supplied by archaeological research. This paper presents the benefits and problems of applying CFD to the hot bath rooms of Roman bathing complexes and provides some preliminary results from the two case studies: a replica bath built for the television series NOVA and the Baths of Caracalla in Rome. The results from the models are extremely promising and demonstrate the importance of buoyancy and stratification: two often neglected properties. Furthermore, the results from the first case study show the importance of doorways to the thermal environment of the room.

Keywords:  
Computational Fluid Dynamics, Thermal Environment, Roman Baths

1. Introduction

This talk describes a synthesis of engineering (more specifically, computational fluid dynamics or CFD) and classical archaeology to develop a more sophisticated understanding of ancient Roman public baths in order to estimate their environment (temperature, humidity, velocity profiles, etc.). Bathing was an integral part of ancient Roman society as there were, according to a catalogue in the 4th century CE (the Notitia Urbis Regionum), nearly 900 baths within Rome itself and many writers at the time reference them in their works. Yet, the surviving primary physical and documentary evidence is sketchy at best. Scholars have studied the architecture of baths and artistic works within but we still do not have a clear understanding of what the environment was like of the heated rooms. Furthermore, these unanswered questions not only limit our understanding of the operation of the bath, but lead to gaps in the social commentary of bathing and the wider Roman society (for instance, cleanliness). Previous attempts to understand this thermal environment have centred on heat fluxes and/or average temperatures and, while important, do not paint an entire picture.

To estimate the thermal environment for the Roman baths, this paper provides two case studies. The first is of the replica baths built by a team led by Fikret Yegül for the American TV series NOVA (Yegül and Couch, 2003), henceforth termed the NOVA baths, and the second is the Baths of Caracalla in Rome (DeLaine, 1997). The NOVA baths proved to be an excellent methodological test for this new application of CFD because the archaeological uncertainty is eliminated as Yegül gives the complete dimensions in his 2003 JRA.
The Use of CFD to Understand Thermal Environments Inside Roman Baths: A Transdisciplinary Approach
Taylor Oetelaar et al.

Furthermore, data collected at the time of construction provide an opportunity for comparison of the results. The Baths of Caracalla, on the other hand, present a chance to apply CFD to a large complex. As the second largest and most complete complex in Rome, its remains are documented, eliminating some of the uncertainty. This paper will centre on the results from the NOVA baths but will also look at the preliminary results from Caracalla and the problems with transitioning between these two.

2. Background

There are three fundamental concepts that are necessary before proceeding further – the general makeup of baths, the heating system, and the basics of CFD. During Roman times, there were two major classes of bath buildings: the balnea and the thermae (Yegül, 1995). The balnea were smaller buildings that were often privately owned whereas the thermae are larger, were state owned, and encompassed other functions. The Baths of Caracalla (Fig. 1) constructed between 212 CE and 216 CE, fall into the latter category while the NOVA baths fall into the former. The main bathing rooms common to most complexes were the frigidarium, the tepidarium, and the caldarium. The frigidarium was a giant basilica-type room that housed the unheated baths. The tepidarium was a smaller room containing the warm baths. The caldarium was another large room that contained hot pools. This project will focus on the caldarium.

The heating system, known as the hypocaust (Fig. 2) consisted of three major components – the praefurnium, the suspensura supported by pilae, and the tubuli. The premise behind the hypocaust was simple but very effective. The praefurnia or furnaces burned fuel that was converted into hot exhaust gases. These exhaust gases were forced between the suspensura or suspended floor and the foundation floor, up through the tubuli or box tubes inlaid in the walls, and out the chimney, all the while giving off radiant heat to the room.

CFD is a branch of fluid mechanics that uses computers to solve problems with no known analytical solution. Since the governing equations for the flow of a fluid, either gas or liquid, are impossible to solve explicitly (i.e., to find functions that can describe the flow fully) except for very simplistic regimes, CFD uses numerical approximations to determine the velocity, temperature, humidity, or other quantities of interest.

In CFD, the first step is to break the volume of fluid down into small, regularly shaped pieces and assign each corner a ‘node’ value, a
process known as meshing. The quality of the mesh dramatically influences the accuracy of the results. Normally, decreasing the spacing between nodes improves the solution accuracy but also increases the computation time and demands. Once the modeller has created the mesh, the geometry is complete. In order to solve the problem two more items of information are necessary: conditions at the spatial boundary and the initial conditions at the beginning of the process. The CFD software then computes the governing equations repeatedly and generates a solution.

Even though CFD is a well established technique, a well-documented case study is highly recommended when starting on a new application. This is where the NOVA baths come in (Fig. 3). They are relatively small, measuring 7.6 metres long by 6.4 metres wide by 2.8 metres high to the base of the vaulting, and has three rooms. The caldarium is in the southwest corner of the building and its air volume is just over 42 cubic metres. The south wall has an alcove for a hot pool. The north, west, and east walls have tubuli up to just below the springing of the vault. Finally, the west wall contains three small windows totalling two square metres.

The applications of computer programs to bathing complexes are few and all led by Basaran (1998, 2005, 2007) but these are all focused on the hypocaust not the room above. There are, however, a number of analytical studies. Thatcher (1956) and Ring (1996) analyzed the Forum Baths at Ostia to determine if glazing was necessary for nude bathing. Jorio (1978-79) and Rook (1978) performed the heat loss studies on the Stabian Baths and the Welwyn Roman bath, respectively. Yegül and Couch (2003) also took measurements of various parameters when they ran their replica – many of which we used in our simulations – and performed a heat loss analysis. All of these studies, though, deal in heat fluxes or average temperatures, which do not provide as detailed an analysis as CFD. CFD has the ability to show how the temperature is distributed and the locations of any drafts and hot spots which could dramatically affect our understanding and interpretation of caldaria.

3. CFD Setup

We carried out the investigation of the thermal environment using the CFD package FLUENT 6.3/ANSYS FLUENT 13.

Since the air volume is quite complex, we started with a more rudimentary mesh to complete the preliminary setup, including evaluation of turbulence and density models and temporal dependency. Turbulence is a measure of the randomness in fluid flow. It is
most recognizable as eddies or white water in a stream. In CFD, there are many different ways of calculating it but the main methodologies are, themselves, an approximation which can influence the results dramatically. For the more complex phenomena, we developed a more efficient mesh which had just over one million cells (Fig. 4). While important, the preliminary setup is beyond the scope of this paper, however, we would like to emphasize the boundary conditions. For the NOVA baths, there are two significant zones: the heated surfaces and the doorway to the *tepidarium*. For the heated surfaces (walls and floor) we chose to use a convective boundary because the hot exhaust gases moving behind supply the heat. Fluent needs a coefficient that relates the heat in the exhaust gases to the heat of the wall, known as a convective heat transfer coefficient (CHTC), which we calculated experimentally. We decided to model the doorway to the *tepidarium* as a neutral configuration of pressure outlet on the top half and a pressure inlet on the bottom half.

### 4. NOVA Results

Displaying meaningful results from the interior of a 3-D model is somewhat difficult. We chose to use planes in the x- and z-dimensions to show the results as they best capture the major features of the NOVA bath. We used contour plots for the temperature distribution and vector plots for the velocity profiles. Contour plots have bands that represent an interval of, in this case, temperature, and the narrower the band, the quicker the change whereas vector plots have arrows that give both magnitude and directional information.

The first thing that one notices with the temperature profiles of the NOVA baths (Figs 5-6) is their non-uniformity and banded nature. This stratification is something that has been neglected in previous heat flux analyses. With CFD, we simulate the entire volume as a moveable fluid instead of a single entity, which
means the hotter air is allowed to rise to the top while the cooler air sinks to the bottom. As such, without a fan to stir the air, buoyancy and stratification have a pronounced impact on the flow. From a comfort perspective, however, this stratification has a drawback. It means that the region occupied by the patrons (i.e., the volume of air below the height of the door) is only between 30°C and 35°C which is only slightly warmer than the next room which is a warm room. This means the region that is occupied by the patrons is necessarily cooler than the average room temperature, though, how great this difference is probably depends on the height of the room.

The next thing that is clearly shown by both the temperature distribution (Fig. 5) and the velocity profile (Fig. 7) in the x-mid-plane is the apparent influence of the doorway. The air flow coming into the room cools down the lower region tremendously which is also the region where the patrons are. This indicates that most of the heat generated by the hypocaust is 'lost' to the unoccupied vault. One can see that the hypocaust is affecting the flow because there is a film of hotter air next to the wall here and along the floor as well as an upward draft along the wall.

A surprising fact while running the simulation was insensitivity of the solution to the CHTC of the tubuli (Figs 5 and 8). The initial simulations of NOVA used an estimate of the CHTC and the experimentally determined value was actually three-hundred and fifty percent of this. With such a dramatic increase in the value, one might expect a corresponding change in the thermal environment of the room; however, there was not. The average temperature in the volume increased only 2°C, from 32.3°C to 34.6°C, which is less than seven percent. This proves that, while the CHTC of the heated walls is a driving force of the simulation, its value has only a moderate impact on the environment which is somewhat surprising as, theoretically, the CHTC is directly proportional to the heat input. The reason for this discrepancy is most
likely that the temperature is dominated by the flow from the doorway which did not change. The extra heat transferred by the change in CHTC could not overcome the cooling effect of the draft from the next room.

In terms of velocities, apart from the influx through the doorway, they are low, almost stagnant; however, there are two smaller drafts that are interesting. One is coming off the windows (Fig. 7) was a result of the cool temperature of the outside in close contact with the window. The more fascinating current, though, is the one from the pool (Fig. 9) and carries water evaporated from the pool. The humidity profile of the z-mid-plane (Fig. 10) echoes the consequence of the pool. There is a definite stream of water vapour rising from the interface and the vapour culminates at the apex of the alcove vault before dispersing into the room. The humidity itself, however, does not have a noticeable effect on the temperature. This is not surprising as most of the room is dominated by the humidity at the tepidarium door boundary condition.

To test the effect of the tepidarium doorway boundary condition, we modelled both the caldarium and the tepidarium. The benefit is that between the tepidarium and frigidarium there is a wooden door that provides an easily-simulated, definite boundary. The addition of the tepidarium allows the free exchange of air that would naturally take place to happen. The initial results for the environment of the caldarium are interesting. Using the same parameters as before, the temperature (Fig. 11) is much higher; however, the distribution pattern is very similar. The velocity vectors from this x-plane are even more intriguing because, although subdued considerably, the exchange between the two rooms is still present, particularly going from the caldarium to the tepidarium. However, these results show that, while the boundary condition we used in the initial runs demonstrates the same basic flow characteristics, it is still not accurate. One improvement might be to change
the temperature of the boundary condition. We had used the average temperature of the tepidarium as given by Yegül and Couch but it might be better to use the average of the two average room temperatures.

There is another aspect that might affect the transfer between these two rooms – a cloth door. In the NOVA special for which this bath was made, people are shown pulling a curtain-like door to after entering the caldarium. The complicated feature about a cloth door, unlike a solid wooden door, is the breathability of fabric, that is, a cloth door has porosity. We modelled a full-length cloth door as well as a door with gaps at the top and bottom. The results from the full cloth door show a definite change from those without. Overall, the caldarium is hotter and there is a visible break in the temperature distribution (Fig. 12). The velocity profile (Fig. 13) is also dramatically affected as the flow out of the caldarium is interrupted creating a recirculation draft at the bottom of the room.

The results from the NOVA baths demonstrate how CFD can be applied to analyze the thermal environment of the caldarium. Even though this building exists and thus can be physically investigated, we can use this methodology to investigate other balnea or small buildings. These findings, more specifically, show how the air stratification means that the coolest air is where the patrons are – a fact which scholars have not discussed before. The other detail that this study brings to light is the impact of doorways and doors to the thermal environment. As the largest air exchange, the doorway allows lower temperatures to enter while letting a portion of the heat out.

5. Transitioning to the Baths of Caracalla

When moving from the NOVA baths to the Baths of Caracalla, however, there are a number of stumbling blocks. The biggest and most daunting is the massive difference in scale, both in the structure itself (Fig. 14) and the actual volume of air in the caldarium (Fig. 15). To quantify this, the volume of air in the caldarium of the NOVA bath is approximately 43 cubic metres or roughly the volume of a semi trailer whereas the caldarium of the Baths of Caracalla is over 45000 cubic metres or about the size of a modern covered stadium. There are four ways of dealing with these issues when discussing the Baths of Caracalla. First, much of the mesh can be coarser as the apex of the dome is over 40m above the floor and the average person is under 2.5m. Second, to use the symmetry of the room and only model a slice of it. While this method is not as accurate as modelling the entire room, the results should capture the major trends and patterns.
of the room and be within an acceptable margin of error. Third, to use multiple runs, each a different slice with a different architectural feature, whether door or window. Together, these will paint a composite picture. Fourth, to use more computing power. This is the most logical but unfortunately most problematic solution in an academic setting due to the limited resources available.

Outside of the computational concerns with making the transition, the incomplete nature of the Baths of Caracalla presents two major problems. First, the tubuli exist presently only to a maximum height of about 71 cm. This is clearly a fraction of their extent in antiquity, but we do not know how high they reached in antiquity. Five metres? Ten metres? To the springing of the dome at 28 metres? As should be clear, this height would have had an impact on how much heat was produced and thus how hot the room became. The best way to investigate this is to do multiple iterations and test varying heights and compare the results. If the tubuli did not reach the dome, there is the dilemma about venting the exhaust gases. We know, however, from the remaining evidence, that the tubuli along the pool could not have reached the roof because of a partition separating the two windows so how were these vented? Second is the absence of glazing materials, which leads to the ongoing debate – were the windows glazed? Like the height of tubuli, this would have had a dramatic influence on the thermal environment of the room. Again, as mentioned earlier, CFD can test various configurations of window glazing and the effects can be observed. The results from simple two-dimensional models...
suggest that, without glazing, the difference between the interior and exterior temperatures is minimal; meaning, in order to maintain high temperatures with such massive windows, the Romans most likely had to employ glazing.

6. Preliminary Results from the Baths of Caracalla

For the initial model of the caldarium of the Baths of Caracalla, we decided to model approximately one-fourteenth of the volume. This is a vertical slice that incorporates half of one window and half of the tubuli that separated the windows. The volume still was just under 3400 cubic metres or approximately one and a half Olympic sized swimming pools which necessitated a coarser mesh than the one in NOVA. Even so, the mesh with 150mm cells has over 2.5 million cells or eight-hundred-thousand more than the NOVA model with the tepidarium.

The initial results from the Baths of Caracalla are intriguing but promising. At the time of the CAA, we have twenty minutes of simulated time. The temperature distribution (Fig. 16a) and velocity profiles (Fig. 16 b) at ten minutes show a updraft in the middle of the room which is perplexing as the heat is primarily coming from the edges. By twenty minutes (Fig. 16 c, d), however, the draft has subsided somewhat which would indicate there is a shift taking place. We must remind you, though, that these results are still very early in the simulation and, therefore, the currents and temperatures are still underdeveloped. We expect the hot region seen near the floor to rise toward the dome leaving a cooler region for the patrons and the velocities to decrease.

7. Drawbacks of CFD

As the results depict, CFD provides a way to predict how a fluid behaves under a certain set of circumstances; however, there are drawbacks. The first and foremost is the experience required to run a simulation such as this as the selection of analysis properties is challenging and can lead to inaccurate results. One of the earliest runs of the Baths of Caracalla, for example, we did completed but the temperature was reaching in excess of 5000°C. We never fully determined what the error was but the anomaly has subsequently disappeared. The second obstruction is the time it takes to run a complete simulation. The final NOVA run of just the caldarium took, using four high-performance processors, three weeks to calculate one hour of actual simulated time and, including the tepidarium, took upwards of two months. In essence, CFD is best for in-depth analyses. The third, especially in a research setting, is the cost of the program. FLUENT can cost anywhere from $10000 to $25000 per year: a price that many researchers cannot afford. There are, however, some free open source codes emerging onto the market, such as OpenFOAM®.

8. Conclusions

In summary, CFD, when applied to ancient Roman baths, can illustrate aspects that are not available through other methodologies. The results from NOVA show that stratification is a key feature of its thermal environment. Despite an average temperature of 34.6°C, which corroborates the data from Yegül and Couch (2003), the temperature profile in Figure 8 shows that the temperature where the patrons are is, for the most part, less than 30°C. This means that the average temperatures used by scholars do not paint a complete picture of the thermal environment. Furthermore, the added level of detail achievable with CFD provides archaeologists a way of determining the hot or cool zones which is not possible with previous methodologies. Another key finding is the relative insignificant impact on the simulation from changing the value of the CHTC: increasing the average temperature by only 2°C. Finally, the importance of doorways
and doors become apparent. The addition of a simple cloth door increases the temperature by a minimum of 5°C. The NOVA results also provide a starting point for future studies of small baths. However, translating this to the Baths of Caracalla is fraught with both computational and archaeological concerns. The early findings from Caracalla are promising but underdeveloped. Overall, CFD can provide archaeologists with another venue to test new ideas and theories.

References


Structural Assessment of Ancient Building Components, the Temple of Artemis at Corfu

Georg Herdt, Aykut Erkal, Dina D’Ayala and Mark Wilson Jones
University of Bath, UK

Abstract:
The pediment of the temple of Artemis at Corfu is one of the very earliest remains of monumental Greek construction. This temple, which dates towards the beginning of the 6th century BC (to judge from the style of the sculptures of the pediment), has suffered extensively since antiquity, leaving us with only a highly fragmented image. Excavated by Gerhardt Rodenwaldt at the beginning of the 20th century, the Artemision was reconstructed on paper by Hans Schleif, in their 2 volume publication. The reconstruction has been criticized in minor respects, but the plausibility of the pedimental sculptures in their proposed position has been generally accepted. In this paper however structural analysis suggests that the current position of these sculptures is inconsistent with stability. This triggers a revised reading of the architectural reconstruction, calling into question the whole plan, with its supposed grand pseudo-dipteral layout.

Keywords: Ancient Greece, Structural Analysis, Finite Element Method, Corfu, Conservation, Architectural Sculpture

1. Introduction

Due to the ruined state of so many ancient structures, the efforts of architects and archaeologists are essential to restitute their original appearance by means of traditional drawings, and, increasingly, using virtual techniques. However, the consideration of whether proposed reconstructions are structurally stable is often not confronted adequately. Unless physical reconstruction (anastylosis) is involved, issues of structural feasibility may not be addressed, or only indirectly by virtue of common sense and reference to comparanda. On account of the structural conservatism of many Greek trabeated buildings this situation is generally unproblematic; yet there are some cases where structural analysis can be usefully employed to validate – or not – a particular reconstruction, even one that has been widely accepted and reproduced in the literature.

2. Temple of Artemis

The Temple of Artemis on the island of Corfu, according to Gottfried Gruben, is the “first and mightiest example of developing stone-architecture” (Gruben 2001, 112). Erected in the early 6th century, it holds a key place in the development of the Doric order that went on to govern the appearance of so much Greek architecture. The pediment of the temple is important as the first surviving substantially complete assembly of Greek architectural sculpture. In addition the monument is striking in terms of its structure and dimensions. As reconstructed by the excavators, it measures 47.90 metres long, 22.41 m wide and about 14 m high, incorporating a complete circuit of columns (peristasis) with 8 on the ends and 17 on the flanks (Fig. 1).

The significance of the Artemision is highlighted by the fact that only a generation or two earlier the great majority of temples in Greek territories were much smaller and narrower; they were built of perishable materials and involved relatively modest spans; their appearance was characterized by ornament and sculpture that by comparison must have been much less elaborate (even if in truth we lack the information to be absolutely certain on this point). In the decades of transition between humble structures of this sort and Artemis’s impressive temple there did exist a few other temples that likewise
deserve to be called monumental. Two of the most notable are the temple of Poseidon at Isthmia (ca. 650 BC) and the temple of Hera at Olympia (ca. 600 BC). The former had a *cella* measuring approximately 32 m by 7.9 m, but it is unclear whether there really was a *peristasis* of wooden columns/posts (Gruben 2001, 106). The Heraion measures 50 m by 18.76 m and is the first building known with certainty to have had a monumental and regular *peristasis*, one consisting of 6 by 16 columns, though initially these were made of wood (Gruben 2001, 52). Soon after our Artemision there followed in Sicily a succession of large religious structures, some of them truly gigantic (Gruben 2001, 288). The Corfu temple was not necessarily the project that launched the new monumentality (this distinction could belong to a different building that happens to be lost to us), but it clearly stands at a critical moment in the development of Greek temple architecture.

The chief focus of this paper is the detailed reconstruction of the pedimented front of the temple and its sculpture as proposed by the excavators. The association of the sculptures with this building is not in doubt; however their capacity to bear load requires evaluation. A revised reconstruction has the advantage of being more plausible from the point of view of statics and constructional soundness, and this in turn has implications for the layout.

### 3. Documentation

Following excavations in 1911 to 1914, the two volume report on the sanctuary of Artemis at Corfu was published by Gerhard Rodenwaldt and Hans Schleif in 1939 and 1940. Their reconstruction of the temple and its dimensions is based on the study of two main aspects; the foundations and the sculpted pediment. Presumably because they were of little use to later builders, the sculptures were not destroyed and remain surprisingly complete. Rodenwaldt and Schleif’s restitution of their original disposition is carefully thought through and represents an important advance in the knowledge of Greek architectural sculpture. The stonework of the foundations on the other hand, was almost completely removed by later builders in the locality (Gruben 1996, 91). On this basis Schleif restores the eight columned front as 22.41 m wide, which is about 3½ m wider than the slightly older temple of Hera at Olympia. The proposed column-spacing is quite dense, with 17 columns reconstructed for the 47.90 m long flanks. Since so little is known about the substructures (the layout of the foundations being determined by the change of fill in the presumed robbed trenches – hardly the most certain method), Schleif sought to give order to the plan by applying a foot grid. The result was questioned in 1943 by Hans Riemann, who suggested a different value for the foot. However, for seventy years no doubts have been raised over the more fundamental
points of Rodenwaldt and Schleif’s work, and in particular the monumental scale of the edifice, the extent of which matches their reconstruction of the pediment (Kähler 1949, 42).

The pediment and its sculptures are indeed impressive. The Gorgon itself in the centre of the tympanon measures more than 3 m in height. The overall length of the pediment is about 20 m, though there is an element of estimation due to some missing parts (Fig. 2). Generally, the sculptural blocks are in pieces but in a good condition; many surfaces are intact, and for some blocks the outline of missing parts can be confidently reconstructed. Each block incorporates a decorated framing strip, one at the top and one at the bottom; the upper follows the pitch of the roof under the raking geison, the other sits on the horizontal geison. The decoration of these two strips is different; a chevron pattern for the upper and a meander pattern about 30 cm tall for the lower.

The sculpted taller blocks at the centre of the pediment present an increased width at the bottom, forming a stand which is preserved up to a maximum depth of about 55.5 cm. The depth of the sculptures is 17 cm at the sides and roughly 29 cm near the centre. This leaves a surplus of the stone material of 26.5 cm to form the background at the bottom. The section steps in higher up, as can be observed at the left panther-lion, and then continues to narrow towards the top, presumably to lighten the blocks for the sake of lifting them (Fig. 3). Around the heads of the two panther-lions the stone has been trimmed down to as little as 5 cm thick behind the feline’s mane. And even worse, the block displaying the Pegasus is reported to narrow down to 4 cm (Schleif 1940, 25).

Four blocks of the raking stone geison remain, each measuring about 2 m in length. The surfaces of these stones show a series of cuttings, allowing a definite allocation of their position and that of the adjoining blocks. At their lower surface an erosion mark (67.5 cm from the front) indicates the junction with the next components. Connections were effected by means of dowels, the cavities for which still remain visible on each geison-block. The upper surfaces of the geison blocks also display a smoothed area as well as holes for connecting to the sima, with dove-tail connectors for joining each block and a large socket at the rear for the timber beams that once spanned across the porch. The smoothed surface matches the dimensions of the second sima, the marble sima, of which some elements survive. This must have substituted an earlier sima, fragments of which also survive made of terracotta and about 80 cm in height, as well as some tiles.

To define the cross-section of the pediment, Schleif placed the sculptures centrally below the raking-geison; their position is set at 67.5 cm from the front by the limit of the erosion due to weathering. The sculptures were thus sheltered from rain, but two problems arise with this position. The first concerns the sequence of construction, since the geison is placed on top of the sculptures. For this to be the case these must have been inserted into position before the roof could have been placed onto the building, the inverse of usual practice. But more importantly: Schleif’s reconstruction presumes that load from the roof can be applied to the sculpted blocks, and he supplies a section that
makes it look plausible (Fig. 4a). However, he does not supply a section through other blocks which have been tapered by the sculptors to the point that their structural capability must be called into doubt (Fig. 4b). Indeed it seems doubtful that Greek sculptors ever intended such a slim profile to bear load. Hence the need to test its feasibility by means of structural analysis.

For this purpose, the weight of the roof has to be assessed. The surviving blocks were combined into two alternative reconstructions by Schleif, of which one results in a lighter construction for the roof than the other. Solution A places the monumental terracotta sima directly onto the massive geison block, while Solution B engages an additional layer in between these two components. Solution B therefore raises the sima about 25 cm, which is consistent with a decorative bead at the edge of the geison. According to Schleif it is the presence of the bead that makes Solution B more probable than Solution A (Schleif 1940, 57 Abb. 43.). However, this fill layer brings additional load onto the sculptures and the slender wall to which they are attached. In order to check the level of potential stresses, structural analysis using Finite Element Modelling has been performed. The presumed properties of the materials and loads, necessary to perform the analysis, are given in Table 1. These have been chosen from the literature (Schleif 1940, 21) on the basis of the archaeological evidence using limit conditions assumptions. In addition to the self-load, tiles have been applied as uniformly distributed loads over the roof and terracotta as a line load on the outer edge of the roof. More accurate results could be obtained by a detailed survey of the remaining fragment and characterisation of the original materials.

4. Structural Analysis

The Finite Element Method (FEM) has been used to structurally analyse the proposed reconstruction of the pediment. The model replicates the left lion-panther sculpture and the objective is to verify whether the thinner portion of the stone section would be able to withstand the roof loads as proposed in Schleif’s hypothesis B. After generating its geometry in AutoCAD, the model has been transferred into the finite element software ALGOR (Fig.

<table>
<thead>
<tr>
<th>Structural Materials</th>
<th>Non-structural Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Properties</strong></td>
<td><strong>Liime Stone</strong> (raking geison)</td>
</tr>
<tr>
<td>Mass Density</td>
<td>3500 kg/m³</td>
</tr>
<tr>
<td>Elasticity Modulus</td>
<td>30 Gpa</td>
</tr>
<tr>
<td>Poisson Ratio</td>
<td>0.25</td>
</tr>
</tbody>
</table>

**Table 1. Material Properties and Loads.**
Meshing (discretization) has been applied, paying attention to each part in the model. As such, the sculpture and the background stone have been finely meshed to better analyze the critical thin section where higher stress concentrations are expected, while the other parts have been assigned coarser meshes (Fig. 6). The model was constructed using 8-node 3-dimensional brick finite elements having 3 translational degrees of freedom at each node. The advantage of using these elements is to provide a complete 3-D state of stress in the stones. Boundary conditions at the base of the stone have been assumed as fixed, i.e. the base of the stone is not able to move downward under the applied load. Although this is not necessarily the case, as some deflection in the bottom of the pediment between columns might have occurred under the weight of the stones above, the error committed is of the same order of magnitude as the other numerical assumptions in the model.
5. Analysis Results and Discussion

The results show that the maximum displacement magnitude is as expected at the edge of the pediment overhang, a very modest 0.156 mm (Fig. 7). The critical stresses are detected at a position just below the strip decorated with the chevron, where the depth of the background stone is at its minimum.

The distribution of the maximum principal stresses shown in figure 8 demonstrates that tensile stresses are generally very low except for a concentration of stress at the necking of the block between the body and the head of the panther where the cross-section reduces. The value in this area is as high as 0.71 MPa. Values of strength for Corfu Limestone have not been found in the literature, though a typical average value for limestone tensile strength is in the range of 1 to 2 MPa. This allows only a very modest safety factor, which would be in contrast with the usual overdesign of Doric temples. Figure 9 shows the distribution of minimum principal stresses, i.e. the value of maximum compression at any point in the stone. In general the values are very low and far from the typical compressive strength of a limestone, usually in the range between 20 and 200 MPa, for un-weathered un-fractured limestone. However it should be noted that where the cross section is reduced, just above the head of the panther, the stress is at its worst, 4.45 MPa. Given the way in which the load is transferred from the geison, performance is best simulated by considering a concentrated load over the thin slab of stone. The strength to point load of limestone, is much lower than its uniaxial compressive strength at about 2 MPa to 10 MPa, hence again creating a condition which would lead to either a very small safety factor or even failure. The results, therefore, suggest that the structural stability of the structure is not fully ensured for Schleif’s hypothesis B.

Thus numerical analysis, thanks to the 3D modelling, is able to identify the concentration of stresses resulting from the specific proposed geometry and construction, and it shows that the
sculptures would not be able to bear the assumed loads. Inclusion of likely additional components not included by Schleif - for example a roof acroterion - would make the situation even more severe. Indeed such a feature might well be expected, and in this period acroteria were large and prominent and therefore heavy. The weight of any such embellishment would arrive at the edge of the verge, increasing stresses on the sculpted blocks underneath. Furthermore, if the designer of the temple intended to bring load upon these elements, he surely would not want to weaken them by tapering. Schleif's solution opposes normal principles of craftsmanship by transmitting structural forces through fragile parts of great value, due to an artistic effort invested in the sculpture.

6. Discussion

As the structural analysis indicates, Schleif's proposal loses credibility due to the risk of failure involving the most precious components of the building, inviting alternative solutions. Indeed, it seems more likely that the sculptures were disconnected from the structural system. This can be achieved by reconstructing the sculptures closer to the front of the pediment and placing a proper load-bearing wall at the back (Fig. 10). This relatively minor change has significant implications. The supposed horizontal geison with an inclined leading edge has to be altered into a more typical flat-topped geison, the common shape for horizontal geisa, which would now host the sculptures up to near its edge. This in fact was the normal condition judging from later well-preserved temples (Schwandner 1985, 108).

Due to the geometry of the corner of the pediment, the modification of a shallower geison results in a more compact elevation, which would be at least 1 m shorter and could perhaps be as much as 3 m shorter. Such a reduction would require rethinking not just the façade but also the plan as a whole. A more compact elevation favours a façade with six instead of eight columns, like so many other temples of that period. A front with seven columns is also hypothetically possible. As part of his eight-column-wide proposal Schleif imagined what the Roman architect-writer Vitruvius called a pseudo-dipteral layout, and both the resulting elevation and the plan have been widely reproduced since their publication seventy years ago. A pseudo-dipteros fits within the overall envelope of a dipteral or double colonnade, except that there are no inner
rows of columns. This involves a wide span for the peristyle; indeed at Corfu – according to Schleif – this reaches almost 5 m, which is to say substantially greater than that of the span of the temple interior. Be it noted that Vitruvius writes that this pseudo-dipteral scheme for temples was only invented by Hermogenes, who worked in the early second century BC. Can we be sure that such an enterprise was possible four centuries earlier, not long after 600 BC? Greater confidence over the design of the Artemision can probably only be achieved by inspecting once more what remains of its foundations. In fact, an archaeological re-evaluation of the site now seems imperative.

7. Conclusions

This paper reveals that the hitherto accepted reconstruction of the Temple of Artemis should be called into doubt. Schleif’s solution involves putting load onto the sculptures and this does not stand up to scrutiny, nor a complete three-dimensional structural analysis using Finite Element modelling. Assessing ancient building components structurally help us find plausible positions for them. It is a tool that can be used to guide new reconstructions and anastylosis, or for scrutinizing published proposals as in this case. By combining structural analysis with attention to the archaeological details of surviving blocks it should be possible to reconstruct ancient monuments in ways that are structurally plausible.

References

Autodesk Algor Simulation Professional. 2011. Autodesk Inc, USA.


Final Results of the Virtual 3D Reconstruction of the East Pediment of the Temple of Zeus at Olympia

András Patay-Horváth
Hungarian Academy of Sciences /University Eötvös Loránd, Hungary

Abstract:
The paper presents the results of a three-years project concerning a major monument of ancient Greek art and approaching a century-old controversy in a new way by producing a virtual 3D reconstruction of a monumental marble group of the 5th cent. B.C. Digital models of the statues were produced by scanning the original fragments and by reconstructing them virtually. The virtual model of the pediment surrounding the sculptures was prepared on the basis of the latest architectural studies and afterwards the reconstructed models were inserted in this frame, in order to test the technical feasibility and aesthetic effects the four possible arrangements. The complete model can effectively be used to verify the results of earlier or more recent reconstructions elaborated with plaster models or presented in simple drawings and leads to the conclusion that the reconstruction most commonly accepted today is actually the least probable one.

Keywords:
Classical Greek Art, Marble Sculpture, 3D Scanning and Modelling, Publication on CD-ROM

1. General Introduction to the Subject

The temple of Zeus at Olympia was built in the first half of the 5th century B.C. (ca. 475–455). Its sculptural decoration consists of two pediments and twelve metopes. Given the large size of the building itself, the sculptures were all well over lifesize and were made of white parian marble. A large number of fragments survived which are conserved in the Archaeological Museum of Olympia and in the Musée du Louvre at Paris. Most of them are quite well preserved and are depicted in practically every handbook on Greek art or on ancient art in general, because nowadays they are generally considered to be one of the most important and most magnificent works of ancient Greek art.

The sculptures of the temple in general and the fragments of the east pediment in particular have been thoroughly studied since their discovery in the 1880’s, but they still pose some important questions, as indicated by the growing number of monographs and scholarly articles related to them (e.g. Treu 1897, Ashmole and Yalouris 1967, Simon 1968, Säflund 1970, Herrmann 1987, Kyrieleis 1997, Barringer 2005, Rehak and Younger 2009). The most recent debate has started with a series of publications (Patay-Horváth 2004; 2005; 2006; 2008) and concerns the interpretation of the east pediment (Fig. 1), which involves the problematic issue of the correct reconstruction of the central group (Fig. 2) as well.

2. The Problem

The arrangement of the five central figures of the east pediment has been the subject of scholarly debates since the discovery of the fragments more than a century ago (Herrmann 1987; Patay-Horváth 2008). The basic problem is that the fragments themselves can be arranged in four substantially different ways (Fig. 3) and there are no obvious clues for choosing the most probable one. There is a fairly detailed description of the group by Pausanias, who saw it in the 2nd cent. AD, but his text (V 10, 6-7) is not conclusive regarding the precise arrangement of the figures (he does not specify how to understand his indications „to the left” and „to the right” of the central figure).
The findplaces are not unequivocal either, since the pieces were scattered around the temple by an earthquake in the 6th cent. AD and the fragments were subsequently reused in medieval buildings.

3. A Brief History of Research

Since the original fragments are insufficient to answer the question and their enormous size and weight make experimentation practically impossible, scholars had to approach the problem in a different way. At the end of the 19th century, plaster models of the statues were produced first on a reduced scale (1:10), then on the actual scale (1:1) and lost body parts, arms, etc. were reconstructed as well. Experimenting with the plaster models for several years, G. Treu the archaeologist, who published the sculptures of Olympia, claimed in 1897 that one of the four conceivable arrangements (Open “A”: K – G – H – I – F; Fig. 4) is physically impossible, because the left hand of figure K and the spear in the right hand of G do not fit but run across each other in the limited space (Treu 1897, 122). To support this rather strong argument, Treu added that with the help of the plaster models, anyone can verify his statement. Indeed, during the following decades, several archaeologists exploited the possibility and experimented with the life-size models: they concluded that the reconstruction proposed by Treu had to be modified at some major points, yet none of them advocated the option excluded by him (Studniczka 1923; Bulle 1939).

The large plaster models (kept in Dresden) were not used for experimentation after the World War II; in fact their sheer existence fell
into oblivion. (It is something of a miracle that they survived the notorious demolition caused by the bombings of the city.) Most scholars used either the reduced models or just simple drawings based on them to propose new reconstructions. Besides a great number of studies, a complete monograph was also published on the east Pediment in 1970, but no one was able to present a fully satisfactory and convincing reconstruction. It is characteristic of the situation that a pair of renowned English-Greek authors presented two completely different reconstructions side by side in the same volume on the sculpture of the temple (Ashmole and Yalouris 1967).

There was a major methodological problem as well. In general, scholars were accustomed to discuss the reconstruction and the interpretation together and the reconstruction was normally adapted to the interpretation, which is logically the wrong way, of course; evidence, which could be used to establish the correct reconstruction independently from the interpretation (i.e. the size and elaboration of the sculptures, which give us a clue about their position in the pediment; the architectural setting of the group; the findplaces of the fragments; the directions provided by Pausanias), was usually neglected.

After a while it seemed that all conceivable arguments have been formulated and no approach proved to be entirely viable, thus archaeologists grew tired of a seemingly unproductive debate and gradually agreed (during the 1970s and 1980s) on a reconstruction (open “A”), which was proposed by a few authoritative scholars supporting their notion by some theoretical considerations of supposed universal validity.

Thus an awkward situation emerged: nowadays, the most widely accepted reconstruction is precisely the one, which was deemed technically impossible by Treu (Fig. 4). Obviously, this would not present a problem, if his results had been thoroughly tested and clearly refuted, i.e. if anyone had showed that Treu had experimented with ill-restored models or had come to wrong conclusions for some other reason. Instead, everyone (except for Grunauer 1981) has ignored his arguments and his results without any discussion. Apparently nobody realized that the best evidence for the benefit of experimenting with life-size models is provided by G. Treu himself, who had advocated the arrangement widely accepted today, while he only had the miniature models at his disposal, but later his experiences with the life-size models made him change his mind.

4. The New Approach to the Problem

Since experimentation with the precious and monumental original fragments is out of question, plaster casts and models are expensive to produce and not easy to modify and to handle, it seemed to be reasonable to apply the latest 3D scanning technology to the problem. The aim of the project was to test the practical feasibility and aesthetic effects of the possible arrangements with 3D models of the reconstructed statues. The digital models were produced by scanning the original fragments and by reconstructing them (i.e. completing their missing limbs and armour) virtually. This enabled the alteration of many details (rotation of the single figures, form and position of the reconstructed parts) and thus the effective testing of every conceivable reconstruction. The technical details and the practical difficulties encountered during the data capture and the process of creating the digital models have been described in previous reports (Patay-Horváth 2010; 2011a).

5. The Final Results of the Project

The experimentation – described in detail earlier (Patay-Horváth 2011b) – with the reconstructed models revealed that contrary to the expectations based on the results of the early experiments with plaster casts, every arrangement could be realized (Fig. 5).
The complete virtual model shows, however, that the arrangement, which was considered to be physically impossible in the 19th century (open “A”) and which is most commonly accepted today, is indeed the most difficult to realize (Fig. 6): the spear of G does not necessarily run across the arm of K, but the distance between them is so small (max. 10 cm) that we can hardly believe that this arrangement could follow the original intentions of the designers or the sculptors.

Furthermore, the model clearly shows that in the case of both open arrangements, another serious problem arises: the spears in the hands of the male figures fit the available space only if both of them grip the shaft directly under the spear-head (Fig. 7), which is otherwise not attested in Greek art.

In the case of closed arrangements, we have no such iconographic problem with the spears (Fig. 8), these arrangements can therefore be regarded more probable than the open ones. Further archaeological considerations (Patay-Horváth 2008) support the hypothesis that the closed “A” arrangement (Fig. 9) should be considered as the most probable reconstruction.

6. New Forms of Presenting the Results

Preliminary results of the project were regularly presented on various meetings and international congresses and were published in due course (Patay-Horváth 2010, Patay 2011a, Patay 2011b), but all these publications (both digital and printed media) were restricted to 2D format and did not enable visualization in 3D. An appropriate documentation of the final results can, however, be conceived only in 3D and the most convenient solution seemed to be the publication of an interactive, multimedia CD-ROM (ISBN 978-963-284-196-0). This final task has also been realized during 2011 and the resulting documentation – already presented at the 23rd CIPA Congress in Prague (cf. Patay-Horváth 2011c) – can be consulted by
anyone wishing to learn more about the details of the project, but for the sake of convenience, a brief description of the main features is included here.

Our goal was to present the 3D models in a fairly good resolution and in a way, which enables the user to manipulate (to rotate, to zoom, to move) them in a relatively easy and uncomplicated fashion, without the need to purchase costly software products (and to learn, how to use them). At the same time, to preserve intellectual property rights, we did not want to disclose the original 3D data captured or created during the project. (They can be obtained on request – mainly for scientific purposes, with no commercial implications – from the author, if both the German Archaeological Institute and the Greek authorities agree.)

Since the project is a multidisciplinary one making use of the latest technological innovations and concentrating on a very specific and complex archaeological problem, it seemed to be reasonable to envisage a mixed audience consisting of both classical archaeologists / students of art history and computer scientists / experts in multimedia visualization. The inclusion of at least some pieces of basic information for both groups was deemed to be essential.

Because the monument investigated during the project, the temple of Zeus and its sculptures are very well-known and famous pieces of the European cultural heritage (the site itself belonging to the UNESCO World Heritage), it was intended to present the project and the models at different levels, not only for specialists, but also for the interested general public.

Our aim was to create a clear and logical structure enabling easy orientation and navigation for every interested party. We chose therefore a format, which combines the appearance of a traditional printed publication with the extended functions of a website. By inserting the CD-ROM into the computer (PC or Mac), the user is automatically confronted with a screen, which functions like an ordinary website with an animated flash intro and a dynamic, multi-level menu (Table of contents) on the left. The content itself is structured in fact like that of a book and the appearance resembles that of a printed book as well (all pages numbered consecutively and having clearly defined dimensions and a constant layout fitting the screen). The pages cannot be scrolled down, but there are arrows on the left and on the right of each, to turn over to the following or to the previous one. In addition there is a navigation bar on top of each page, directly below the title. By clicking on this, a complete scrollable list of all pages (with their individual titles) appears on the screen and the user can easily move to any other page, he is interested in (Fig.10).

The text contains links to attached documents of various kinds (e.g. scholarly publications in pdf for specialists, reports in mp3 and avi format for the general public) and to other pages of the book guiding or informing the user, like cross-references and footnotes of a traditional book. Images and 3D models displayed on the pages can be enlarged and viewed in a separate window by clicking on them. In order to ensure wide and easy usability, 3D models were included in 3D pdf format. This enables the user to observe the
models from any point of view and to enlarge any part of them, but the original 3D data sets are not disclosed (Pletinckx 2011).

The fragments of the pedimental figures have been generally designated by alphabetic letters since their discovery and precisely because their arrangement in the pediment is disputed, they were arranged in alphabetical order, one figure per page. Navigation between them is facilitated for the non-specialists by a page showing miniature icons of the models and the commonly used designations of the figures, both functioning as a direct link to the page, where the models of that particular figure are displayed. On these pages, the model on the left shows the surface of the preserved torso as recorded by the 3D scanner, the one in the centre displays a closed digital model of the piece, whereas each one on the right presents the whole figure as completed during the project, the original parts displayed in grey, the completed ones in pale blue. Textures taken from the present state of the fragments were not applied to the models, because they are irrelevant for the project and because they are generally misleading, since ancient marbles were originally colored in general, and in this case practically every trace of polychromy has completely disappeared (Brinkmann and Primavesi 2003).

The four different virtual 3D reconstructions of the central part of the pediment are displayed in a similar way (the original and the completed parts differentiated by the same colors and with a navigation aid showing all variants side by side). Two pages are devoted to every single arrangement showing the model from three different but constant viewpoints (all of them on the main axis of the pediment): 1. “museum view”: viewer standing approximately on the same level as the statues (as in Fig. 5); 2. “ancient view”: viewer standing approximately on the ancient ground level before the temple (Fig. 11); 3. “aerial view”: from above, pedimental frame removed from above the statues (Fig. 12). In addition, by clicking on the museum view, each possible arrangement of the central group can be viewed and manipulated in 3D pdf format. With the help of these models, everyone can decide which option seems most or least satisfying technically and aesthetically. The most probable reconstruction of the entire pediment (according to the author) is also included and can be studied in 3D pdf.

Texts, presentations and audio-recordings of lectures, interviews of various genres are displayed in unaltered form (each one of them in the original language, i.e. English, German, Hungarian or French). The differences are due to the various types of audiences (specialists or general public) and reflect at the same time the progress of the research. Published and forthcoming manuscripts of the author are also included in the appropriate sections.

Numerous photographs of each pedimental figure are also added in the Gallery section and may thus be compared with the 3D models. The aesthetic value of these images cannot be denied, but at the same time, they
clearly show the limitations of this kind of traditional documentation.

A comparison with the published audiovisual documentation of similar projects (e.g. the Trier Constantine, the Metopes of Selinunte or the virtual reconstruction of Olympia by the Powerhouse Museum, Sidney) was already presented by the present author (Patay-Horváth 2011c) and shows the conceptual novelties and technological improvements introduced in the structure, design and content of the CD-ROM.

A publication of the results along with the 3D models produced during the course of the project is also intended in the first volume of the newly established on-line journal “Digital Applications to Archaeology and Cultural Heritage” (DAACH), which offers the unique possibility of publishing interactive 3D models along with related articles.

7. Conclusions

The complete virtual 3D reconstruction of the composition leads to the conclusion that the reconstruction, which is most widely accepted today (Fig. 4), is technically the most difficult to realize and that both open arrangements would be feasible only if we ignored a general pictorial convention of ancient Greek art. Still, it is important to emphasize that the virtual reconstruction does not enable us to establish the right arrangement, i.e. the one actually realized in antiquity, but only to exclude (with a high degree of probability) two of the four options. However, considering the uncertainties experienced so far, this result can be regarded as a great progress. Though the remaining two closed arrangements are possible both technically and iconographically, one can observe, that every piece of evidence, which is independent from the interpretation actually points to type “A”, which can be considered therefore as the most probable reconstruction (Fig. 9).

The project has thus reached its major goal and contributed significantly to a debate, which engaged archaeological research for more than a century. It demonstrated at the same time, that 3D technology can be used not merely for documentation (as it is most frequently employed), but for effective research purposes as well.

References


283. Remshalden: Greiner.


Teaching Cultural Heritage and 3D Modelling through a Virtual Reconstruction of a Medieval Charterhouse

Andres Bustillo, Ines Miguel, Lena Saladina Iglesias and Ana Maria Peña
University of Burgos, Spain

Abstract:
We present an educational experience conducted with high-school students, introducing them to the generation of 3D computer graphics and to virtual reality environments applied to cultural heritage. The duration of the experience was 40 minutes. It was intended to convey to the student an understanding of the basic concepts of 3D computer graphics and the main historical and artistic characteristics of a medieval building: the Charterhouse of Miraflores (Burgos, Spain). During the experience, students could modify the 3D model of the historical building, its textures, and its lighting, and interact with the 3D model in a Virtual Reality Room. The assessment of this experience was based on surveys filled in by the students. The survey results suggest that the proposed methodology is an effective way of transferring different types of knowledge and of stimulating the interest of students in various disciplines, such as art, history of art and 3D graphics.

Keywords:
Virtual Reality, 3D Modelling, Teaching, Cultural Heritage

1. Introduction

Computer Graphics has become a powerful tool for entertainment in our society. Unfortunately, high-school students are very often unfamiliar with the process of generating 3D computer graphics, hindering professional development in this discipline. An initial step would be the development of suitable methodologies for short educational experiences, to improve their knowledge of this process, which would serve as introductions to various topics. A second step would be the development of new approaches to teach history and art that could capture the attention of students, in this case for both high-school and undergraduate students. These new approaches should focus on increasing student participation in the learning process and on incorporating more Information Technologies (ITs) in the process, as students clearly pay more attention nowadays, if the learning process includes these technologies (Korakakis et al. 2006).

There are very few examples of specific methodologies for these topics. Some proposals have focused on software tools that could be used in an introduction to the process of generating 3D Computer Graphics (Van Langeveld and Kessler, 2010, Wang et al. 2008) and to Virtual Reality (Roussou, 2004). In other cases, researchers focused on the use of ITs, especially 3D modelling tools and virtual reality environments, for the teaching of topics such as architecture (Styliadis et al. 2008), history (De Paolis et al. 2010, Di Blas and Poggi, 2008) and archaeology (Lucet 2009, Baeza and Barrera 2010). But all these proposals rely on the involvement of students in complex 3D environments, a task that is only possible if the educational experience lasts for a reasonable length of time. Student involvement is not easily achieved through short workshops, which attempt to give a general overview of the 3D Computer Graphics Generation Process or serve to introduce students to the main characteristics of an historical building. The efficiency of short educational experiences is hindered by a lack of trust between teacher and student. Virtual Environments have the potential to counteract
this lack of trust, by stimulating the interest of students and assisting their concentration on specific information.

This study proposes a new methodology for short workshops that responds to both points: it stimulates the interest of students in historical buildings and introduces them to 3D Computer Graphics. The experience was first organized during the tenth Science Week of the University of Burgos, at a workshop called “Virtual Reality applied to historical heritage” (12-19th November 2010). The objective of the Science Week is to bring the research work conducted at universities closer to the public, and especially to high-school students. The intention behind the workshop was to allow high-school students to identify the main stages and concepts involved in 3D computer graphics generation, while also learning about history, medieval thought, and the artistic qualities of the Charterhouse of Miraflores (Burgos, Spain), built in the Late Middle Ages.

The proposed methodology involved four steps. First, two teachers -one from the Department of History of Art and another from the Department of Computer Science- presented the story behind the virtual reconstruction of the medieval church. Then, the students, closely accompanied by the teacher, were allowed to work on the 3D model of the church, shaders and lighting. In a third step, they were allowed to interact, in very small groups, with the 3D model in a Virtual Reality Room. Finally the students were asked to fill in a survey to provide feedback and conclusions on this practical experience.

The paper is structured as follows: Section 2 describes the hardware and software used in the workshop. Section 3 describes the structure of the workshop in detail: goals, steps and the student profiles. Section 4 explains the students’ responses to the surveys, analyzing the extent to which the goals of the workshop had been achieved. Finally, the conclusions and future work are summarized in Section 5.

2. Hardware and Software

The University of Burgos has recently created a Digital Innovation Center, the main purpose of which is to teach students ITs for Digital Media. This Center is composed of a 3D-Graphics Editing Room, a Visualization Room, and a Server Room.

The multimedia content is created in the 3D-Graphics Editing Room. This room (Fig. 1) includes 25 workstations equipped with software for 3D modelling and animation (both Blender and Maya). The workstations incorporate a quad core processor with 2 GB RAM DDR2- ECC HDD memory.

The Visualization Room (Fig. 2) is a semi-immersive 3D virtual reality room. It consists of two 3D projectors, a high-quality sound system, a motion capture system, a 3D positioning system, and a workstation to generate 3D renders in real time.

The Server Room includes a render farm and three high-performance servers: a storage server, a streaming server, and a web server for dissemination of content through Internet. The render farm is composed of an 8-node cluster with a quad core processor, each with 8 GB RAM memory.

Finally, a 3D model of the Church of the Charterhouse of Miraflores, as it was at the beginning of the XVI Century, was also used in the workshop. This virtual reconstruction was possible thanks to the close collaboration
between computer scientists, historians and visual communicators. The 3D model was created using Blender software and has been described in a previous publication (Bustillo et al. 2010). A five-step process was followed to build this virtual reconstruction: 3D modelling, texturing, lighting, rendering and postproduction. These steps have been widely used to build virtual recreations of historical and artistic buildings such as the Cathedral of Halle in Germany (König, 2006) or the Tower of Hercules in La Coruña in Spain (Noya, 2010) instead of using more accurate laser-scanning techniques (Martin et al. 2010). A new modelling methodology, based on simplified 3D models, and a multilayer texture application, was developed to allow real-time virtual recreation of the historical building (Bustillo et al. 2010). Once the 3D model was finished, it was exported to the virtual reality room to recreate a semi-immersive environment of the inner part of this building.

3. Workshop Description

The design of the workshop was complex, due to the multidisciplinary nature of the knowledge to be presented to the students (History, Art and Computer Graphics) and differences in the student profiles. The duration of the workshop was approximately 40 minutes and each group had between 15-20 students. The workshop was held over 6 days with 3 sessions per day and a total of 313 students in attendance.

This section describes the goals of the workshop, the different kinds of students in attendance, the structure of the workshop, especially designed for the proposed goals and the student profiles in the three knowledge areas: History, Art and Computer Graphics.

3.1 Workshop Goals

A workshop should impart knowledge of a varied nature (History, Art and Computer Graphics) and stimulate student interest in this knowledge, to bring students closer to Bachelor studies in these topics. The following specific objectives were proposed to reach those two general goals:

• Understand the various stages in a Virtual Reconstruction of an historical building;
• Assimilate the main historical and artistic aspects of the Church of the Charterhouse of Miraflores;
• Assimilate the concepts of 3D modelling, texturing, lighting and rendering;
• Understand the importance of collaboration between experts in different disciplines to develop Virtual Reconstructions of Cultural Heritage;
• Stimulate the interest of students in Art and History through the use of ITs;
• Stimulate the interest of students in 3D Computer Graphics and Virtual Reality.

3.2 Student Profiles

The students were divided into two different groups in accordance with their main interests. The first group was formed of high-school students studying technical subjects. The activities of the first group in the workshop focused mainly on the process of 3D computer graphics generation and, to a lesser extent, on the study of the historical and artistic value of the building. The second group was formed of high-
school students studying humanities and first-year students following a Humanities degree course. The activities of this group focused mainly on the historical and artistic value of the building and, to a lesser extent, on the process of 3D computer graphics generation.

The gender division, age and numbers of students in these groups were as follows:

- 189 students studying Science Baccalaureate from different high-schools in the city and the province of Burgos. The group consisted of 85 male and 104 female pupils aged between 15 and 21 years, of whom 148 were in the first year of the Baccalaureate course and 41 in the second year.

- 124 students studying a Humanities Baccalaureate from different high schools in the city and province of Burgos. The group consisted of 48 male and 76 female students aged between 15 and 18 years, of whom 112 were in the first year of the Baccalaureate course and 12 in the second year.

3.3 Workshop Steps

The duration of each workshop was approximately 40 minutes with 15 to 20 students in attendance. The workshop was divided (Fig. 3) into four steps: a presentation of the virtual reconstruction of the building, a supervised practical exercise on the 3D Model, an interactive tour through the virtual reconstruction, and a short evaluation of the workshop by means of a survey.

The first step was an introduction, to clarify the main concepts, to the Virtual Reconstruction of Artistic Heritage using the example of the virtual reconstruction of the Church of the Charterhouse of Miraflores. This 7-minute presentation, with slides and a beamer, displayed the main sources to study historical buildings, from historical documents and printings to actual scans, which are needed to create the virtual reconstruction. Then the presentation focused on the construction of the 3D model: the basic concepts of 3D modelling, basic operators (extrude, spin, etc), the differences between procedural and image textures, the different types of virtual lightings, and the concept and computational requirements of rendering 3D computer graphics. The presentation also remarked on the necessary collaboration between computer scientists, historians and visual communicators, to carry out the virtual reconstruction.

After this presentation, the students were given the opportunity in two supervised exercises to modify some elements for themselves of the 3D model of the Church of the Charterhouse of Miraflores. These exercises took around 20 minutes. Just before the exercises, a brief presentation of the Blender interface (movement through the virtual space, selection of orthogonal and perspective views, and selection of objects) was made. Then, in the first exercise, the students had to change the image texture of one piece of stained glass for an image they had downloaded from internet.
(Fig. 4). After finishing this task, the second exercise consisted of modifying the render of the lectern located in the centre of the church (Fig. 5), working on the lighting properties (color, intensity and halo effects) and their positions. In both cases the students were able to create final renders (before and after their modification of the original 3D Model) to evaluate their work. To speed up the render process, the render farm is used instead of the computational capabilities of the workstations controlled by the students. This point was mentioned on by the teacher, although the students had no choice in the matter.

Although the groups were small, with between 15 and 20 students, two teachers are necessary at this stage of the workshop, to achieve satisfactory results within a short time (20 minutes) and to avoid student frustration. One teacher explained the exercises, with the help of a beamer, and slowly worked through the exercises, while the students repeated by themselves what they saw. The second teacher mingled with the students, answering their questions and helping them, if lost. As the time for adaptation to the 3D environment of the students is very different depending on the profile and skills of each student, the most advanced students were able to do the two exercises quickly, and they therefore had the opportunity to move through the different layers that make up the complete 3D model of the church.

After the supervised practice, the students went to the Visualization Room in small groups of 4-5. Once inside, they performed a semi-immersive virtual tour of the Church of the Charterhouse of Miraflores. Before doing so, a teacher introduced students to the components of a virtual environment: 3D beamers, positioning system and 3D glasses. Then, during the virtual tour, the students handled various interactive elements of the 3D model (videos, images, text, etc). At the same time, they discussed the significance of the main elements of the building with the teacher, from the perspective of the medieval builders, and the artistic value of each element (history, style, patrons, etc). This part of the workshop lasted approximately 7 minutes.

In the last step of the workshop, the students had around 3-5 minutes to complete a survey designed to evaluate their level of understanding and their assimilation of the concepts explained to them and to measure their interest in the Virtual Reconstruction of Cultural Heritage. The structure of the survey is described in detail in the next section.

4. Evaluation of the Workshop

This section describes the way the evaluation of the workshop was performed and the results that were obtained. The students
were presented with a survey after attending the workshop that had been designed to satisfy requirements opposing requirements. They were expected to fill in the survey on a voluntary basis and to indicate clear, concise responses in various areas: student motivation, acquired knowledge, and so on. In consequence, a short survey was designed that could be quickly filled in, all the more so as a survey that would have taken longer to fill in than the short educational experience itself would have made little sense.

4.1 Survey Structure

The survey completed by the students after attending the workshop consisted of 13 questions. Eight questions had five possible answers, from which the student had to choose two. Four questions had only two possible answers (yes/no) and the last question was rated on a scale of 1 to 10.

The questions were split into three groups, to evaluate whether all of the workshop goals had been fulfilled. The first group was composed of six questions on general concepts related to history, artistic value and 3D computer graphics; these questions were about the sources of information needed for a virtual reconstruction, the kings that promoted the building, the prominent artists who worked on it, why it is necessary to apply textures on the 3D model and what a render is. The second group included four questions on the advantages and the limitations of virtual reconstructions; the questions referred to problems of real-time rendering and how to display realities no longer available in the real world. Finally, the third group was composed of three questions that evaluated student interest in the real church and in the 3D computer graphics generation process; these questions related to their willingness to visit the real building and to prepare a 3D Model by themselves.

4.2 Results

A total of 226 completed surveys were collected from the 313 students that attended the workshops. The responses were normalized to the number of valid surveys for analysis and representation and the following results were obtained reached.

The students showed an understanding of the main stages of the virtual reconstruction: from the research needed to define the Cultural Heritage building for virtual reconstruction to the Computer Graphics process itself (3D Modelling, shading, lighting and rendering). The responses to the questions intended to evaluate knowledge of the historical and artistic value of the Cultural Heritage building showed that the workshop had helped students develop an understanding of the artistic style of the building, the monarchs that patronized the church building, the principal artists that worked in it, and the century when it was built. Figure 6 shows that more than 80% of students answered these questions correctly in all cases.

When asked about the difficulty of the 3D modelling process, all the students considered it to be complex (Fig. 7). In a range from 1 (easy) to
10 (very complex), the students gave an average value of 7. When gender was considered in the analysis of the answers, women found the 3D Modelling process more complex than men, by a difference of half a point (Fig. 6). The Student t-test was performed, to confirm whether there was a significant difference between genders. The test returned a low level of confidence, leading to the conclusion that this difference was not significant. The same conclusion can also be drawn from the high Standard Deviation in the answers to this question for both genders: 2.24 for woman and 1.72 for men and the small difference in the mean value between genders of 0.5.

Most students, when asked whether they would like to do 3D modelling themselves, gave a positive response (Fig. 8). It is interesting to note the high response to that question, in view of the previous answers of the students, affirming that the 3D graphics generation process was a complex process.

Different gender patterns in behavior were also remarkable in relation to the willingness of students to create a 3D model by themselves (Fig. 9). Men were more interested (82%) than women (56%) in creating a 3D model, although the difference in the perception of the complexity of 3D Computer Graphics generation between genders was not significant (0.5 in a scale of 10).

Most of the students (83%) were more interested in the generation of 3D computer
Teaching Cultural Heritage and 3D Modelling through a Virtual Reconstruction of a Medieval Charterhouse
Andres Bustillo et al.

5. Conclusions

We have described an educational introduction to the historical and artistic significance of a medieval religious building and to fundamental concepts of Computer Graphics. Although of a very limited duration, the 40 minute experience showed how the use of a virtual example and the practical engagement of the student could help students achieve complex objectives. This experience was implemented during the Tenth Week of Science at the University of Burgos, in the “Virtual Reality applied to historical heritage” Workshop. The 3D model was a reconstruction of a building of historical interest: the Charterhouse of Miraflores in Burgos. The results of this practical experience were based on an analysis of surveys completed by the students who participated in the experience.

The surveys evaluated the strengths and weaknesses of this teaching methodology, in addition to levels of student understanding of historical and artistic meaning and the process of virtual recreation. Various clear conclusions may be extracted, the most important of which is that the students, with backgrounds in both sciences and humanities, showed high levels of understanding of the main concepts related to history, art and computer science. They understood the meaning of the building as a holistic representation of the religious and social reality of the Late Middle-Ages. They demonstrated an understanding of the advantages and disadvantages of using 3D Virtual Reality for such reconstructions. These results show how useful a 3D Reconstruction is for non-specialized groups. There is only one gender difference that can be extracted from the surveys: although all the students find the process of 3D modelling to be quite complex regardless of gender, when asked whether they would have an interest in creating such a 3D model by themselves, the male students showed greater interest than the female students. As a final conclusion, it may be said that most of the students after the virtual tour expressed a heightened interest in visiting the building as it is nowadays. These responses show the potential of short educational experiences, not only to teach concepts of a different nature, but also to enhance the interest of students with no specialist interest in Cultural Heritage.
The survey and the organization of this teaching experience reveals some weaknesses in the teaching methodology. The first refers to the technical and human resources and the time needed to perform an educational experience of this kind. The requirements for a 3D immersive environment and the limitation of 5 students in this room per group, although allowing close interaction between students, teacher and the 3D model, also suggests that these resources should be available to medium-size groups of students (20-30 students) for longer periods of time. Also the surveys show that the understanding of basic historical and artistic concepts of the building is not completely achieved; this result begs a further question: could a 20-minute visit to the real church in a medium-size group conducted by a teacher using “edutainment” techniques be better for this task?. This question has yet to be resolved.

Finally, it should be noted that the results of this educational experience have not been compared with similar teaching experiences based on traditional methodologies, perhaps because of the difficulties of using such methodologies to introduce students to relatively recent and varied concepts such as 3D Modelling, History and Virtual Reconstruction. A set of short-term experiences focused on practical exercises might obtain good results in teaching 3D Modelling concepts, although the simultaneous introduction of historical topics would perhaps be an insurmountable challenge for traditional methodologies.

Future work will focus on generating similar methodologies for undergraduate students of engineering to increase their interest in 3D Modelling and Animation, which represent two areas of specialist employment that are expected to grow in Spain. Another new aim will be to lengthen the duration of the workshop, to broaden knowledge of the building’s medieval history and art that could be discussed and taught to the students and to introduce some concepts and related techniques of 3D Animation to students.

Methodological improvements would also be of interest, so that the 3D Model and the multimedia material embedded in it could partially replace the presence of the teacher in the 3D environment, thereby reducing the human resources needed to perform this kind of a teaching experience. Finally, the students that participate in these experiences could be followed up to evaluate long-term motivation and background knowledge.

Acknowledgments

This work has been partially supported by the Program “Impulso de la Industria de Contenidos Digitales desde las Universidades” of the Spanish Ministry of Industry, Tourism and Commerce and the “Aula para La Creación y Difusión de Contenidos Multimedia” of the University of Burgos. The authors would especially like to thank Mrs. Laura Martínez for her collaboration during the performance of the educational experience.

References


3D Reconstruction in Archaeological Analysis of Medieval Settlements

Daniele Ferdani  
National Research Council of Rome, Italy

Giovanna Bianchi  
University of Siena, Italy

Abstract:  
The road inland from Massa Marittima (Italy) takes you to a small town, Montieri. In the Middle Ages, thanks to its metal-bearing deposits, this area was the focus for a complex history, and the town still boasts a range of impressive architectural features, remains of its important past. Using the results of archaeological surveys and historical sources, a complete three-dimensional reconstruction of the architecture of this town was carried out, focusing attention on the most important buildings. The 3D reconstruction helped researchers to make hypotheses regarding the original architecture and layout, and will also help tourists to visualize the ancient site, and understand it better. This paper discusses the project’s methodology, tools, and aims. Special attention is given to the various stages of the work that is under way, and on the types of data used to obtain the 3D reconstructions.

Keywords:  
3D Modelling, Medieval Archaeology, Medieval Architecture, 3D Reconstruction, Photogrammetry

1. Introduction

Given the large number of important monuments and archaeological sites in Italy, many of them tend to receive less consideration than they deserve, or are ignored, even, in favour of better-known sites and monuments. Moreover, many remains are in such a poor state of conservation that only experts or scholars can understand their context.

This is the case with Montieri, a small medieval settlement situated in southern Tuscany, between Siena and Volterra. Although the settlement is not well known, it has interesting examples of architecture, typical of the 12th and 13th centuries. The site has typical residential buildings (“tower-houses”), but, importantly, it also has a mint. This is the mint of the Bishop of Volterra, which struck silver coins. The Bishop’s residence can also be seen, an elite, high-status building adorned with rare architectural elements.

Over the last few years, the University of Siena, in collaboration with local authorities, has been involved in studying the site’s architecture, using photogrammetric techniques and the methodological tools of the Archaeology of Architecture. Within this discipline, which was developed in the context of Italian medieval archaeology, methodological tools have been developed which borrow the principles of stratigraphical sequencing from traditional archaeology, with the aim of identifying, in the fabric of walls, the traces and relationships of the main actions involving construction and destruction which have taken place over time. This allows us to identify the different phases in the life of the building itself, thanks to the identification of the physical and stratigraphical relationships between the various transformations in the construction of the architecture.

To improve and promote knowledge and understanding, the 3D reconstruction images have been used in scientific journals and magazines, and in video form in a documentary, shown on national Italian TV, on the town’s history. The case-study is ongoing, and will be
updated over the next few years as new finds come to light.

2. The Site and the Project

The idea of reconstructing the medieval settlement of Montieri arises from the necessity of reconstructing the most important medieval architectural buildings, built between the 12th and 13th centuries.

In the Middle Ages, Montieri was a castle town governed by the Bishop of Volterra. The town used to operate its own mine-workings, and processed the ore that was mined there. Almost all the architectural features of the centre of Montieri bear signs of the town’s important past. Some of these buildings have retained much of their original appearance, such as the Cassero (the castle “keep”), the fortified headquarters adorned with elaborate sculpted elements, perhaps belonging to the Bishop, who controlled the town. The Biageschi and Narducci towers rise above the old town centre. Although these towers are shorter than their original height, they are wonderful examples of private fortified residences, characterized by the stylistic and architectural influence of Siena and Volterra, and they reflect the economic wealth of their inhabitants. As revealed by our archaeological analysis, other buildings do not retain their original appearance, but some surviving parts, eg part of the town’s outer walls, allow us to imagine how they appeared in the past. This is the case with a large building, which once housed the local mint, now known as the “Foundries”. (Bianchi 2010; Arangouren et Alii 2008). This building consists of a ground floor and an upper
floor. The ground floor is well preserved, and had 9 rooms with vaulted ceilings. The facade consists of stone, with pilasters, and arches made of decorative brickwork, still in part visible. During excavation inside the ground floor rooms, several braziers and furnaces for smelting metal were discovered, along with small channels cut into the ground for mixing the molten metal. The Bishop’s mint was active throughout the 13th century, but the building was radically transformed at the beginning of the 15th century, as were most parts of the medieval buildings.

This is our starting point, in order to re-establish the original appearance of the buildings. From the beginning, the project of reconstructing medieval Montieri followed two related goals: scientific research, and communications.

Scientific research took into account all the surveys and traceable data on the city, analyzing stratigraphical sequencing. This data was processed, and compared with other sources of available historical and geographical information. Indeed, previous experience showed that 3D visualization, and the possibility of virtual reconstruction, can be an additional benefit for discussion, and for decisions regarding analysis and interpretation.

Interpretation was more focused on the virtual representation of the building for the purpose of cultural dissemination. In detail, all the ancient architecture was studied and reconstructed using new 3D technologies, and attention was focused on different fields of activities and cultural diffusion (Coralini, Vecchietti 2007).

3. **Sequence of Work**

The sequence of work, applied to obtain reliable reconstructive models, set out from archaeological analyses to obtain 3D modelling processes, as follows:

The first step was a survey of the remains using the method of the Archaeology of Architecture and metrics, and photogrammetric survey techniques, also relying on the findings of the archaeological excavation conducted in
previous years by the University of Siena. The survey data was supplemented with historical sources from written sources and iconography, land registry maps, and drawings.

The second step was to process the survey data, and to draw up raster and vector-based documentation, in the form of plans, front elevations and sections, essential for archaeological and architectural documentation, and for the subsequent step: Interpretation.

In line with the processed archaeological data and historical sources, carefully matched with survey and photogrammetric data, it was possible to interpret the history of the buildings. Subsequently, reliable and scientifically acceptable hypotheses were proposed, and a set of 3D models, visually representing the original construction phases, were created (Hynst et al. 2001; Borra 2000).

4. Architectural Survey

As a compulsory strategy for surveying the medieval architecture of Montieri, we took advantage of some of the most common approaches adopted in the field of the historical building industry (Medri 2003). In order to get a set of data suitable for archaeological analysis and 3D reconstruction, we had to:

- provide a thorough metric documentation, for precise archaeological stratigraphical analysis;
- compile background records for a complete catalogue of different kinds of stone and stone-work;
- exploit geometric data for 3D modelling;
- acquire high-quality photos for documentation and fine detail;

In order to satisfy these requirements, a special survey was conducted, using different methodologies that are closely linked with each other:

The survey of the buildings was a major challenge, because the buildings were so close together that it was impossible to measure many of their architectural elements.

Thus, in order to measure the facades of such buildings, indirect methods were applied, such as Laser Total Station and Photogrammetry, while a traditional direct survey was used for the smaller interior spaces (laser distance meter, measuring tape).

The first step was the establishment of a global topographical network that constituted the basis for the subsequent external photogrammetric operations, and internal direct surveys, in order to obtain building plans. The topographical network was drawn up using laser total stations, while the photogrammetric survey was conducted on the visible facades of the historical buildings.

Images were gathered from different positions, and using different focal lengths. In this way, inaccessible zones were minimized, thereby reducing a typical fault associated with this approach.

The procedure is summarized as follows:

- Each individual façade was subdivided into small squares on a virtual grid. On each node of the grid, a control point (CP) was placed. These CPs were then measured with the total station, to derive the local coordinates. Finally, for each square on the virtual grid, a photo was taken.
- The post-processing operations were performed by digicad3D, a 3D CAD photogrammetry program, to obtain elevations, drawings, or maps from photos. The elimination of distortions was based on the CPs, captured as reference data. Making use of the photos and the measured control
points, the software processed new, scale images, without distortions of perspective.

• Finally, the adjusted photos were worked into a mosaic, in order to obtain high-quality images of the facades of the surveyed building. This technique, in these particular, difficult conditions, was very well-suited for the survey and analysis of the buildings. Stereo-photogrammetry was not necessary, because the surfaces of the facades are very flat, and do not have features that protrude (such as balconies or decorative elements).

This approach ensured good definition of the facing stonework, by allowing a detailed characterization of each single stone, essential for mapping surface operations, a material characterization map, and stratigraphic profiles.

These high-quality, measured images were very useful for detecting and analyzing the construction phases, and the techniques used for building the walls, for evaluating the skills and capability of the builders (master-builder, bricklayers, stone-cutters), and drawing up hypotheses about the technical environment.

5. 3D Reconstruction

After the preliminary survey, and once all the sources were compiled, the 3D modelling process began. The 3D reconstruction and representation of the buildings was the hardest phase in the research.

To achieve a reliable reconstruction of Montieri and its buildings, several types of archaeological and architectural data, with the support of experts from Siena University, were investigated and interpreted. Firstly, we organized all the information regarding the archaeological remains (layouts, surveys, pictures of the excavations, pictures of the

Figure 5. Rectified photo-mosaics of the “Biageschi” and the “Narducci” towers.

Figure 6. Data post-processing: stratigraphic profile and construction phases of the “Narducci” Tower.
buildings before and after WW2, and materials documentation). Subsequently, we integrated the technical data with thematic data, such as previous publications, bibliographic resources, hypothetical scientific-based reconstructions, drawings, etc.

In order to identify the original parts of each building, and define what they must have looked like in the past, the suggested reconstruction was drawn up comparing the historical, archaeological, survey, and photogrammetric data processed during analysis, and considering the constructive and aesthetical rules of that period (12th-13th century), in order to refine our understanding of the ancient architectural elements, and associated fixtures and fittings.

The hypotheses, when not backed up by actual historical or archaeological sources, were founded on comparative data, or supported by formal rules, construction techniques or medieval modules, identified in other buildings and nearby medieval towns, such as Siena, Volterra, Campiglia Marittima and Grosseto.

The logical reconstruction criteria were adopted as follows:

- **Reconstruction by “analogy”:** The reconstruction is based on analogy with a well-known and recognizable theoretical model. Despite having only one part of an object, the reconstruction was carried out by referring to a widespread standard (e.g., stone-working techniques, and the artistic style of the windows and the doors).

- **Reconstruction by “comparisons”:** The reconstruction is not based on a theoretical approach, but on direct comparisons with extant remains in the local area (e.g., the towers of Montieri were compared with the towers of Siena and Volterra).

- **Reconstruction by “deduction”:** although some buildings or architectural elements were incomplete, their complete appearance is deduced by referring to the formal characteristics of the buildings, or to repeated patterns (e.g., the partial arcades of the Foundries were completed by repeating the pattern of the remaining pillars and arches of the porch).

- **Reconstruction by “hypothesis”:** This is the most complex process. Hypotheses are based on conjecture, or archaeological evidence, (e.g., the wooden elements, like balconies or roofs, were deduced from putlog holes) (Medri 2003).
After the interpretative studies, to obtain the 3D reconstructive models, we start from vectors and the rectified photos as references. Autodesk AutoCAD software was used to extract the principal construction lines from the elevations, to integrate missing parts, and to create orthographic projections. The computer graphics-based reconstruction was developed using Autodesk 3DStudio Max. The 3D modelling work was tackled by using imported 2D plans and profiles. In accordance with the architectural and structural features of the building (all the data used in the reconstruction is directly derived from surveys), the construction of the three-dimensional model maintained the architectural irregularities and asymmetries in the alignment of the walls. All the models were unwrapped and mapped with textures devised “ad hoc”, using an “in situ” photographic campaign in order to make them as real and as plausible as possible. Finally, to create a realistic natural illumination, we used the “daylight system” tool provided by V-Ray, a plug-in of 3DStudio Max.

The first 3D drafts were used as a common basis for discussion and analysis of interpretative decisions, and to refine the reconstruction interpretation. Suggested reconstructions were crucial in verifying the choices that were considered, and to reject those that proved wrong, as our past experience demonstrates. The 3D models were not only the end result of the work, but also a scientific instrument for interrogating the medieval architecture, and understanding its original shape (for example, on viewing the first reconstructions of the Foundries, researchers became aware that the second floor of the building must have had bigger windows, in order to ventilate the rooms where the coins were minted, and thus dissipate the high temperatures created by the furnaces).

The 3D models were also used to analyze the load–bearing structures, and to distinguish between differing phenomena. This allowed us to restore the original, correct floor plan, and make informed inferences about the kind of ceiling (i.e.: tunnel vault, and wooden floor), and the shape of the roofs themselves. Finally, for the excavated buildings suggestions were put forward for the reconstructed sections and interior spaces, and, by referring to archaeological finds, furniture, equipment and tools were reconstructed and located in their original positions.

Up until this point, we obtained the following models as they must have appeared in the 13th century:

![Figure 8. The medieval town of Montieri as we suppose it was between 12th and 13th century.](image)
5 detailed models: 2 residential towers, the Foundries, the church of S. Giacomo, and the so-called Cassero (the multi-apse church of St. Niccolò, is still being excavated).

1 suggested model of the town, referring to the interpreted plan of the medieval settlement (the layout of the historical center has remained much the same).

As mentioned above, the aim of this project was to obtain 3D reconstructions of the main medieval buildings, both for scientific purposes and for cultural dissemination.

The archaeological studies and the reconstructive hypotheses were published in the visitor-guide “Archaeological Landscapes of Montieri / Paesaggi Archeologici di Montieri” and in the hiking-map “Medieval Itineraries of Montieri / Percorsi archeologici di Montieri”. In these publications all the research project is summarized, and translated into language suitable for the “general public”, focusing on the site between the 12th and 13th centuries, narrating the past life, and reconstructing what the panorama could have been in the medieval period.

Moreover, the models were used to shoot architectural walk-throughs, and to film a short movie. Still photos were shown on an instalment of “Geo&Geo”, a scientific Italian TV programme, and shown on national public TV (RAI) to promote tourism in this area, which is

Figure 9. 3D model of the Foundries in the 13th century phase.

Figure 10. Cross section of the Foundries. In evidence: the wooden infrastructure of the second floor and the roof, the blacksmith’s forges and the brazier.
less well-known than others, but which was the focus for a complex history which can still be read today in its remains which are very well preserved in its main centre.

6. Conclusions

This procedure of analyzing and interpreting Montieri’s medieval architecture, and communicating research findings to specialists and to the general public, represents an important model which will be adopted also as archaeological research in this area continues. In particular, a similar procedure will be pursued for the reconstruction of landscapes around Montieri itself, and especially in the case of an ecclesiastical complex called the “Canonica [Church] di S. Niccolò”, situated a few kilometers from Montieri, near the silver mines themselves, and which has been excavated since 2009.

The complex was totally redesigned in the 12th century, with the construction of a church with six apses, the only example of its kind in Tuscany, and a series of buildings adjacent to the church, in part designed for the processing of metals. The buildings, and their richly ornamented architecture, are clues as to the importance of the site, including in symbolic terms, since it was an expression of the power of the Bishop of Volterra, and his ability to control the local, ore-rich area.

The 3D modelling will thus be used to arrive at a full understanding of an entire historical context, not just an individual monument. This understanding will also serve to boost visitor numbers in a zone which is currently situated in an area of marginal interest to tourists.

Acknowledgements

We wish to thank the “Colline Metallifere” Park, the “Soprintendenza Archeologica della Toscana”, the Mayor of Montieri, and the whole team at the LAAUM laboratory (Siena University).

References


Handling Transparency in 3D Reconstructed Online Environments: Aquae Patavinae VR Case Study

Daniele Ferdani, Bruno Fanini, Guido Lucci Baldassari, Ivana Cerato, Sofia Pescarin
CNR-ITABC-Rome, Italy

Abstract:
This paper presents the different steps of a virtual archaeology project where different data are produced using separate pipelines, and then are optimized and integrated in a real time environment, accessible online to specialists (restricted access to back end) and to public (open access to front end). The work is based on an open source tool (OSG4WEB) previously developed by the team that we recently optimized by adding new features and functionalities to enable the final publication of large spatial scale landscape reconstructions and virtual environments. The Aquae Patavinae project was chosen as a case study in order to explain the problematic aspects typical to virtual archaeology work where the main goals are the interpretation and reconstruction of an archaeological landscape through online virtual archaeology tools and making “transparent” to non-expert users the metadata used to construct and ensure the accuracy of the 3D models.

Keywords:
Transparency, 3D On-Line, 3D Reconstruction, Virtual Reality, User Evaluation

1. Introduction

When we approach the problem of archaeological landscape reconstruction, we often have to deal with palimpsests of data, both historical and archaeological, complex and many-sided historical contexts that are difficult to interpret even by researchers. The case study analyses interpretation and transparency issues is the Aquae Patavinae VR. In this project we are investigating and reconstructing the thermal landscape of the Euganean Hills around Montegrotto Terme (Padova, Italy) in cooperation with the University of Padova and funded by MIUR and Veneto Region. In this region, the known sites are spread across a wide area: only one main archaeological site is open to the public, others are still under excavation or cannot be visited, and many others consist only of small little scattered evidences identified through archaeological surveys, historical studies and remote sensing.

It is evident that we are facing a well-known scientific problem, common to archaeology, which regards the interpretation and reconstruction of an archaeological landscape as a single global entity (Barceló et al. 2000). This issue is more relevant in the case of a project that uses online virtual archaeology tools addressed to both researchers and non-experts. There are two aspects of the problem one directed to the “user: researcher” and the second one directed to the “user: visitor”.

In our case study, we have dealt with this issue by adopting two approaches that have taken into account the London Charter (Principles 3 and 4: LC 2009) and Seville Principles. They are:

- development of an appropriate workflow, whose goal is to build reality-based models acquired in the field or through remote sensing technique (e.g. terrain models), and non-reality-based models that include interpretative/simulation models connected to the reality-based ones.
- New release development of a web plug-in for 3D online exploration, “OSG4WEB” (based on OpenSceneGraph, Calori et al. 2009).
In the new release we have added features such as:

- Customization of the navigation system (it changes in accordance with the type of exploration, scale of the landscape and visualization device: fly, walk, touchscreen, natural interaction);

- Development of specific UI (User Interface) which allows user to upload “interpretation models” on the landscape (i.e. GIS-based 3D models visible in transparency above the original 3d archaeological remains) or “reconstructed models” (visible as solid models with texture);

- Real-time access to metadata, containing short descriptions, data sources and level of reliability;

- Ongoing development of tools for interactive vegetation management inside the scene;

- Renewed tools within 3D models optimization for online exploration.

The innovation of this work regards how we faced “transparency” issues from the digital creation of 3D to the online visualization and interaction. Section “Transparency of the data” is specifically dedicated to the explanation of how we have solved this issue. This section is anticipated by a “workflow” chapter that describes the methodological approach which allows us to achieve these goals, manage transparency issues and address implementation details about OSG4WEB plug-in and its exposed user interface.

2. Workflow

2.1 From Fieldwork to 3D Reconstruction

The reconstruction and web 3D publication of the Montegrotto thermal landscape has been an important goal and challenge for the “Aquae Patavinae” project. In order to rebuild the entire city and territory, with all its buildings and monuments, different approaches have been adopted:

Active and passive scanning technologies to survey the archaeological remains and to get the reality-based models;

3D modelling (procedural and Computer Graphic) to build the modern city and the reconstructed archaeological monuments (the missing architectural parts), bearing in mind the need to optimize the 3D models for online use (e.g. low-poly models, squared “atlas” textures, texture baking, etc.)

Surveying Archaeological Sites

At the moment, there is no single reality-based modelling technique for all types of objects and sites of high geometric accuracy that is able to satisfy all requirements: portability, full automation, photo-realism and low cost, flexibility and efficiency. The two methodologies, range and image-based modelling (Remondino and El Hakim, 2006, 269-291), are able to satisfy some of these requirements, and together they ensure the best results in terms of reliability and completeness. For example, in order to acquire highly-detailed metric data on archaeological sites (e.g. a Roman theatre; a hydrological system with channels and baths; a Roman villa), a geometrical survey, based on range and image-based techniques, was conducted. Due to the presence of modern protective coverings, surveying of the buildings turned out to be very difficult and problematic. Many parts of the archaeological structures were hidden and the entire direct access was impossible. Thus, in order to obtain a successful survey and also to try out various techniques, two 3D acquisition methods were used: a) Photogrammetry; b) Laser-scanner. The use of these two different methodologies turned out to be appropriate for geometric control of the objects, comparing the actual shape to the original design, and achieving photo-realistic texturing.
3D Modelling

Two different non-reality based approaches were used to reconstruct the Virtual City of Montegrotto:

**Procedural Modelling** was employed to model the modern city (as it is today) (Parish and Muller, 2001, 301-308). As for improving realism and contextualize excavated sites, a virtual reconstruction of approximately 6000 buildings in the modern city was created. The use of CityEngine software allowed the overall entities in the hierarchy of the scene to be controlled and a large number of diversified, schematic and low-polygon buildings to be obtained. However, to create different building types the various typologies of them needed to reproduce a reliable virtual copy of the city, as modelling it house by house proven to be challenging. We used shape grammar rules to define variables, based on random criteria. Moreover, in order to determine the external appearance of each building (the details come from texture not geometry), a correlation matrix that randomly assigns a texture was developed: into that matrix, dimension information is associated with each texture so that, when the software generates the buildings, it automatically gives the correct height and proportion.

**Computer Graphic Modelling** was employed to design archaeological features. The reconstruction of the Roman monuments was carried out starting from scanner laser and photogrammetric models (used as reference for the low-polygons modelling) taking advantage of a “hand modelling technique”. The goal of these 3d models was to obtain a Real-Time (RT) hierarchical version of each of them that could be explored online. To model them, Blender and 3DStudio Max software, were used. Each monument was subdivided in different levels of details, using a simple constraint such as the final dimension of each block. We chose a polygonal modelling approach to reconstruct the architectural features because it is very flexible and appropriate to create several versions of a model with different numbers of polygons. Regarding the optimization of the models for real-time exploration, each model was further decimated in the modelling software, subdivided in small portions and exported by material in .obj format. The final step was to create and apply squared “atlas” textures with color, ambient occlusion and bump information.

3. Data Transparency

Similar to the majority of the complex and multidisciplinary studies, our project involved managing a wide variety of information and data. Normally these data remain hidden to users (Forte and Pescarin 2009) who thus have a vague perception of the entire workflow, which leaves them with no idea of the metadata used to construct and ensure the accuracy of the models.

Although the new 3d technologies used for disseminating cultural heritage allowed us
to create a more efficient and clear image of the ancient world, there were still the following problems/questions to be overcome/resolved:

- Why should data used in the reconstruction processes really need to be transparent?

- How to handle and improve the transparency, reliability and uncertainty of the data in our project?

- How can virtual reality help us improving the dissemination of cultural heritage information while make people aware of the past in a more effective and interactive way?

Through these questions, in the Aquae Patavinae project, we tried to overcome the problem of the transparency, offering users an online virtual reality navigation system with different data layers and tools permitting them more in-depth exploration of the archaeological information.

3.1 Virtual Reality Navigation System Based on OSG4WEB

The main aim of our project was a virtual real-time online exploration system. Therefore, we concentrated our efforts on managing and publishing a large archaeological landscape, 3D models and metadata on the web while simultaneously enhancing the system’s reliability.

As a matter of fact, the virtual reality system approach helped us to create a web platform based on the open source tool (OSG4WEB) for sharing archaeological data, 3D models and interpretations, and manage reliability through the user’s interaction by transforming their passive role to an active one (Forte and Pescarin 2009, 198). Details about OSG4WEB implementation are as follow

- Within 3D visualization and remote content fruition inside a web browser, there is actually no absolute standard defined (Akenine-Möller, Haines, 2002, 477-479). During the development of such components, able to display within a web page a complex virtual world, the main objectives are:
  - Rendering: visual quality and overall performance in large environments
  - Network/World wide web: latency reduction and workload on 3D data streaming
  - Deployment: how easily/realistically can the system be deployed in the present web reality

In order to achieve these objectives and to provide a rewarding experience during the real-time exploration of the virtual environment, the first release of the plug-in was further developed and enhanced. The plug-in is based on the open-source framework OpenSceneGraph (Kuehne B. and Martz P., 2007), which suits...
our objectives and gives special attention to a multi-resolution approach, exploration, interaction and online use.

The system provides a FrontEnd and a BackEnd, both web-based and running on a simple LAMP server: a restricted team can access the BackEnd using an assigned username and password, with full access to management of the virtual world scene graph, 3D models and interface elements. The whole interactive scene is then displayed using the FrontEnd, a component used by the general public, that can be inserted in any web page, with fully customizable options. The navigation system provides enhanced walk and fly modes, aided exploration within hotspots or areas of interest, customizable constraints and some basic physics effects, such as gravity, inertial forces and surface frictions. Depending on the actual scenario, users can modify the main interface with specific icons, a compass and other 3D elements: each element can be customized. The plug-in addresses web and online data streaming issues by providing advanced segmentation, different levels of details, multi-resolution and paging techniques. The “Aquae Patavinae” case required particularly the online management of a large territorial model, procedural models and complex 3D models. This management needed a set of batch tools to rearrange, optimize and organize the 3D content and its hierarchy. To generate the actual terrain model the tool “osgDem” was used to create an organized hierarchy and paged tiles with on-demand loading as the user approaches a certain spot, providing an overall well-balanced web load.

### 3.2 Transparency in the Virtual Reality Navigation System

During the VR navigation, information is provided to users. Three information layers can be visualized as users move around the virtual world, offering the possibility of exploring the archaeological/contemporary landscape (actual model), interpreted landscape (interpreted model) and reconstructed landscape (reconstructed model). These three “landscapes” allow users to take advantage of different point of view on the past:

- **Actual model**: this layer shows the archaeological sites as they appear today: the modern city and environment and the archaeological sites. The models visualized with this option come from reality based modelling
- **Interpreted model**: it shows the archaeological site as they appear today but placed above the ruins transparent 3D models of monuments as we suppose (based on archaeological/historical evidence) they appeared during Roman times. In this way the users can make comparisons to better understand the remains.
- **Reconstructed model**: shows the monuments as they potentially were during
Roman times in solid appearance, i.e. 3D models are not transparent.

During the exploration, essential information about the buildings and their architectural parts are provided using informative floating panels (Fig. 3). Thus, there is continuity between the virtual tour and data regarding the models; the user can visualize both without stops. These panels are activated only when the user is in proximity to hotspot position and provide information on:

- Name and function of the building;
- Level of certainty;
- Typology of buildings and elements;
- Approximate dating;
- Main references used for reconstruction and bibliographic sources.

In order to show and make clear to the user the different levels of certainty concerning the reconstructed architectural models and preserve the realistic manner of the models (Forte 2007), we avoid using colour coding (Jones 2009, 102), instead the information about certainty, is presented in an encoded format that takes advantage of stars. This system is simple and straightforward:

* one star means “evocative”: no archaeological information available, the reconstruction was based on comparisons and specific criteria (usually referred to decorations, paintings, wooden elements).

** two stars mean “possible”: little archaeological information, the reconstruction was based on the evidence and comparisons (usually referred to supposed architectural elements of which there are many gaps nevertheless their shape is deducible - due to repetitive pattern - such as colonnade, roofs, gardens).

*** three stars means “probable”: a lot of archaeological evidence, the reconstruction is based on well preserved architecture, such as, a theatre, that maintains its original plan and part of the height.

Trying to reconstruct the archaeological landscape and the potential ancient landscape means facing many “gaps”, areas with little or no information. We tried to overcome this problem by allowing the user to explore in real-time the entire landscape, enabling him to freely explore the entire landscape, both at a territorial scale (macro: fly-mode) and at a site-scale (micro: walk-mode). To do this, two kinds of exploration modes (fly-mode and walk-mode) and some aided navigation tools (snapping spheres and path-constraints), were developed.

The “Fly-mode” option allows users to explore the landscape. In order to keep the user from “landing” outside archaeological areas, a
“height constraint” was set at 300 meters height. In this way we prevented the user from getting lost in the city or from being in blank-area where reconstructions have not been implemented due to the lack of data. Archaeological areas that can be explored are pointed out using “snapping spheres”, transparent bubbles, placed above the sites. When these spheres are clicked, users are brought inside an archaeological site and the software changes the navigation mode to the “Walk-mode” option. When the “Walk-mode” option is enabled, users can move around the archaeological area, making a virtual tour of the ruins and the reconstructed architectures. At the same time, however, users are unable to leave the sites to visit the surroundings. As mentioned above, using this approach users are constrained/guided along predetermined paths, avoiding the gaps, but they are free to choose the order of the visit, in the meanwhile the archaeological sites and monuments.

4. Output, Type of Interactions and User Evaluation

During the development of the application, different interaction modalities were considered: three kinds of interactions were developed for different types of users:

• Online desktop-based VR exploration for home users: the visualization and exploration of the virtual world by users within a webpage is available through a plug-in installed in a common web browser (www.aquaepatavinae.it, in the virtual reality section)

• Touch screen interaction for on-site small exhibition: a kind of natural interaction which belongs to the daily life of almost every potential user

• Camera Tracking interaction for on-site exhibitions: a natural interaction using low cost camera (a webcam) to track movements of visitors within a defined field. This kind of interaction may be programmed in the future for use by home users with disabilities.

Our objective is to provide users with a welcome screen of an overall view of the actual landscape (Montegrotto city), a brief instruction panel about navigation and an interface to enter and explore a few hotspots of archaeological sites. These hotspots are scattered across the landscape and invite the visitor to enter and interact (mouse-click or tap on touch screen devices) with archaeological sites. The generic walk-through session, once user has entered a specific site, consists in activating different phases or “virtual layers”, starting from archaeological landscape (the archaeological remains as they are visible today) to interpretation (remains with transparent 3D reconstructions) and final reconstruction (full solid 3D reconstructions with modern city fading away). The interpretation layer particularly allows visitors to make a comparison.
between ancient and modern times, while the reconstruction phase lets users explore a more realistic depiction of ancient Roman times (for example with steam simulations in thermal baths or “piscine”). The user controls this timeline, moving from archaeological to interpretation to reconstruction phases, by switching side tabs (mouse and touch-screen interface) or by passing through visible “time portals” (only natural interaction interfaces).

The testing protocol used to evaluate Aquae Patavinae application within natural interface outputs was developed within the V-MUST.NET framework (V-MUST 2012) and influenced by previous results and observation sessions carried out during public exhibitions, in particular ArcheoVirtual 2011 (Archeovirtual 2011) exhibition. Later on, at the Montegrotto public event, a renewed evaluation protocol was used, in particular, new evaluation protocols were conceived to identify and measure specific variables (visibility, mapping, consistency, user feedback and constraints) and to create a protocol able to spot usability bottlenecks and other issues. In order to measure the quality and usability of the interaction systems, observation sessions were carried out to:

- Measure duration of interaction,
- Determine which interface elements were mostly used to access content,
- Calculate time required by visitors to familiarize with a specific natural interaction and how this improves their learning benefits,
- Evaluate comprehension of icons and time portals,
- Evaluate comprehension of informative floating panels,
- Record user comments and behaviors in certain situations,
- Record any technical issues (if any).

Collected results, showed that the majority of users were very involved in the session, with a positive attitude towards experimenting new interaction systems and exploring the virtual world contents of “Aquae Patavinae”. The biggest bottleneck that emerged regarded the guidance of the visitors during the initial stage of the of body tracking interaction (a team member always had to initiate and supervise the visitor).

This situation is presently being investigate to overcome the problem. The ultimate goal of these ongoing evaluations (observations, surveys and direct interviews) is to understand what impact particular “transparency” methods and metaphors have on user experience and understanding. In the future, the results of these evaluations will be used to modify and improve our current approach.

References


3D Documentation for the Assessment of Underwater Archaeological Remains

Barbara Davidde Petriaggi, Roberto Petriaggi
Istituto Superiore per la Conservazione e il Restauro, Rome, Italy

Gabriele Gomes de Ayala
Naumacos, Italy

Abstract:
In 2001 the Superior Institute for Conservation and Restoration (Istituto Superiore per la Conservazione e il Restauro) launched Restoring Underwater. Restoring Underwater is a project aimed to the study and experimentation of instruments, materials, methodologies and techniques for the restoration and conservation in situ of ancient submerged artefacts. Set up in 2001 it started by restoring the vivaria of the Roman villa of Torre Astura (Nettuno- Rome). Since 2003 the research has been focused on the submerged archaeological site of Baiae where over the years the restoration of sectors of certain buildings within protected marine area were carried out: the so called Villa con Ingresso a Protiro, the Villa dei Pisoni, the so called Herculanea street and the Building with porticoed courtyard near Portus Iulius. Recently a new type of relief was carried out to document the phases of the restoration of a room paved with opus sectile, situated in the Underwater Park of Baiae not far from the Nymphaeum of Punta Epitaffio (-5 m) using Naumacos L1 scanner laser. Naumacos L1 system has been created specifically for the archaeological research and for the stratigraphic survey during the underwater excavation. The L1 system generates a cloud of dots to create a photographic textured model, that is accurate within a millimeter. L1 system can scan big areas and automatically merge them into a mosaic, getting to an improved sub millimeter level of precision, which means that it’s possible to obtain an archaeological survey of smaller details. The tridimensional laser scanning survey plays a prominent role in the planning and finished stage of the restoration in order to show the restored piece in a museum. This method of 3D documentation shows better the state of conservation of the monuments, and increase the value of scientific dissemination.

Keywords:
3D Documentation, Underwater Archaeology, Underwater Restoration, Baiae, San Pietro in Bevagna Wreck

1. Introduction

The 3D documentation project is part of the project Restoring Underwater launched in 2001 by Istituto Superiore per la Conservazione e il Restauro (Superior Institute for Conservation and Restoration) (Petriaggi and Davidde 2007; Petriaggi and Davidde 2008).

Restoring Underwater, in accordance with the Convention on the Protection of the Underwater Cultural Heritage, Unesco (Paris, 2 November 2001), is aimed to the study and experimentation of instruments, materials, methodologies and techniques for the restoration, documentation and conservation in situ of ancient submerged artefacts.

Set up in 2001, this project started by restoring of three basins of the vivaria of the Roman villa of Torre Astura (Nettuno- Rome). Since 2003 the research has been focused on the submerged archaeological site of Baiae (Petriaggi and Mancinelli 2004).

Baiae was a famous seaside town much prized in Antiquity for its temperate climate, beautiful setting and the properties of its
mineral waters. Since the second century BC this area was the most popular resort of Roman Aristocracy and the Imperial Family up to the end of the third – IV century AD, when the bradyséisme caused its submersion. This archaeological site offers a wide variety of architectural structures from luxurious maritime villas and imperial buildings to modest residences, private thermae, tabernae and other structures typical of a Roman city.

The Underwater Park of Baiae was created on August 7th 2002 by the Ministry for the Environment, acting together with the Ministry of Cultural Resources, Transport and Agrarian Policies, and with the Regione Campania. The Soprintendenza Speciale per i Beni Archeologici di Napoli e Pompei has held provisional responsibility for its management. The submerged area of about 176.6 hectares safeguards the archaeological remains of the Roman town of Baiae and the infrastructures of Portus Iulius and represents an underwater area of great environmental value.

At Baiae over the years we carried out the restoration of some areas of certain buildings: a room paved with a white mosaic and a calidarium of the Villa con Ingresso a Protiro, a sector of the Villa dei Pisoni, a portion of the so-called via Herculanea and a part of the republican Building near Portus Iulius (Petriaggi 2005). In 2007, 2009, 2010 three new targets have been added to the research:

- Nine cast iron cannons discovered offshore the coast of the Marettimo Island (Sicily, Trapani). Here we have applied for the first time in the Mediterranean, a sacrificial anode to limit the effects of electrochemical corrosion for the in situ conservation (Bartuli et al. 2008).
- The roman wreck carrying a load of sarcophagi off the coast of San Pietro in Bevagna (Taranto). Here, to test an appropriate way to present to the public the underwater archaeological remains, we have tried illustrative panels encased in stainless steel boxes. The project also offered the opportunity to test the Scuba Scan system to obtain a three-dimensional survey of the sarcophagi.
- A traditional fishing boat recently discovered off the coast of the Martana Island (Bolsena Lake, Viterbo). Here we realised a system to reinforce the wreck planking by using plexiglass and polycarbonate bands.

2. The Underwater Archaeology Project and 3D Documentation: San Pietro in Bevagna Wreck and the Underwater Site of Roman Baiae

The systems available to researchers and archaeologists to detect an underwater archaeological 3D area, was essentially, the multi-beam systems and the interpretation of stereoscopic pictures. The first system produces a 3D graphic representation of the seabed morphology, particularly suitable to the geological study of the conformity of the sea-beds and underwater canyons. The second system produces a 3D model through the interpretation of stereoscopic pictures via a particular software created for this kind of use (available freeware) that interprets the pictures and as a result it obtains a 3D model roughly like the original one. (Canciani, Gambogi et al. 2002; Mahon, Pizarro et al. 2011; Diamanti, Georgopoulos et al. 2011; Gearhart 2011).

The underwater archaeological remains of a wreck carrying a load of sarcophagi, which
had foundered off the coast of San Pietro in Bevagna, covers an area of circa 148 sq m. The site, approximately 100 m from the coast and 4-6 m in depth, contains twenty - two partially worked white marble sarcophagi of diverse shapes and sizes, each weighing between 1000 and 6000 kg.

Based on the type of sarcophagi and the characteristics of the ceramics embarked, as well as on the comparison with the contemporary wreck at Methoni (Greece), the date proposed for the shipwreck is around the first half of the 3rd century AD. From the study of the data that can be deduced from the extension of the site and layout of the load, the San Pietro in Bevagna ship may be about 20-22 m in length and 5-6 m in breadth, but so far no surviving elements from the wooden structure have emerged.

Probably the sarcophagi were bound for Rome, where they would be unloaded at the Statio Marmorum of Ostia; from here they would be transported on the naves Caudicariae by river to the Ripa Marmorata, close to Monte Testaccio, and then to Campo Marzio, where the marble-workers plied their trade at the Statio Rationis Marmorum.

As said above this archaeological site offered us the opportunity to test a system to obtain a three-dimensional survey: the ScubaScan System. This system was planned and developed by Menci Software in collaboration with the Suor Orsola Benincasa University of Naples and ASA Studio (Petriaggi and Davide 2010).

The instrument used to conduct the survey consists of three calibrated digital cameras, attached to an aluminium bar and positioned in parallel. It measures $80 \times 38 \times 22.5$ cm and weighs 10 kg. The cameras are housed in watertight coverings and linked in such a way as to allow synchronized shots, making it possible to obtain images of the same place or the same object from three coordinated points, the precise distances of which (i.e. distances between one camera and another) are known. This condition is necessary for the subsequent software processing that transforms the raster data acquired into the corresponding three-dimensional models. The instrument undergoes two types of calibration: first of all, a one-off calibration of the cameras, carried out at the Menci laboratories, determines the optical features of each camera and lens. The system is then re-calibrated each time the instrument is dismantled and re-assembled, so as to reconsider the alignment and position relative to the cameras on completion of assembly.

The cameras are set so as to optimize the quality of the image according to the prevailing conditions of light and water transparency. The standard working distance for the proposed layout is between 1.5 and 6 metres.

The photographic shots acquired underwater are downloaded and processed within the PC.

The software consists of a specialization of the ZScan software that is able to:

- put the synchronized images from the three different cameras in order;
- calibrate the system based on the appropriate shots at the calibration polygon;
- permit the user to insert the salinity levels recorded underwater;
- define, for every triplet of images comprising one shot, the area useful to reconstruct,
as well as the final resolution of the model obtained;

• generate the 3D model of each shot;

• triangulate these models, so as to obtain a mesh onto which a raster texture is automatically applicable. The triangulation allows the subsequent automatic generation of accessory products such as orthophotos, dem (Digital Elevation Model), elevations, level curves, etc. within the software of Menci Software, Z-Map. The work of assembling the obtained models, implemented by integrated software modules again developed by Menci Software, then makes it possible to obtain an overall three-dimensional model. This latter procedure, depending on the characteristics of the images or the subject, can be both automatic and manual.

For the 3D survey of the sarcophagi from the wreck of San Pietro in Bevagna forty-eight single models were produced and later assembled in one overall model for the survey of the sarcophagi. The resolution of the final model obtained is on average 1.5 mm between one three-dimensional point and another.

During the acquisition phase of the images it emerged that certain external factors, almost always difficult to control, such as the variability of the light conditions under the water, the presence of fish and suspended algae, or the existence of groups of artefacts very distant from one another, can create problems in producing the 3D survey, substantially altering the texture information legible at the same point on an artefact recorded at two different moments in time.

In order to minimize these inconveniences, we realised it is advisable to adopt certain measures: as far as possible, the photographs should be taken at a time of day when the light is greater and as constant as possible, avoiding, for example, cloudy days; if the underwater site contains substantial suspended matter or animal life, and also when groups of artefacts are very distant from one another, it was noted that the 3D scanning was assisted by positioning a topographic net with appropriate reference marks over the site, before proceeding with the photographic stage. In this way it is possible to avoid all the problems relating to the discontinuities of the surface.

The experimentation has shown that, if used following the above-mentioned methods, the ScubaScan system is very effective and makes it possible to reduce the time and lower the costs of underwater surveying considerably. Moreover, the 3D models obtained have the characteristic of being geometrically and chromatically faithful to the reconstructed object, as well as being exportable in various forms and usable in numerous contexts.

On June 2011, a new type of relief was carried out to document the phases of the restoration of a room paved with opus sectile (Fig. 1), situated in the Underwater Park of Baiae not far from the Nymphaeum of Punta Epitaffio (-5 m) using Naumacos L1 scanner laser3.

Naumacos L1 is the name given to the first ever 3D underwater scan laser projected, developed and realized by the researcher

3 The archaeological and topographical study of this room and its relation with the thermal bath located near the nymphaeum and the petrological study are in progress.
Gabriele Gomez de Ayala to be used in aquatic environments (Gomez de Ayala 2008).

Naumacos L1 system has been created specifically for the archaeological research and for the stratigraphic survey during the underwater excavation. Uncountable issues were solved to use the laser technology, adapted to scan underwater archaeological sites. One of the main issues solved is the refraction.

One of the main goals, while realizing Naumacos L1 system, was the one to offer an instrument able to make a precise and scientific survey, feasible in time, way and formality that the aquatic environment and its different and changing conditions impose on archaeologists.

Naumacos L1 is an underwater 3D scan system light and compact (15 x 16 x 10 cm - 4 kg) based on the emission of a laser ray that hits a point and to that point are associated some coordinates in the 3D space; the addition of more points generates a 3D model of what has been framed, obtaining technically a cloud of points; simultaneously Naumacos L1 samples the texture of what is being scanned, that is the picture with precise coordinates that allows the system to perfectly project the same on scanned surface, obtaining a photo realistic model.

Picturing to survey an archaeological site, the need will be the one to make a series of 3D scanning to be joined one to the other, respecting their own position obtaining just one model of the site; that model allows the archaeologist and researcher to make every kind of measurement: height, width, depth of each single detail, distance and angles among the single objects from there proceeding to the creation of excavation and topographic maps.

The L1 system generates a cloud of dots to create a photographic textured model, that is accurate within a millimeter. L1 system can scan big areas and automatically merge them into a mosaic, getting to an improved sub millimetric level of precision, which means that it’s possible to obtain an archaeological survey of smaller details (Figs 2-3).

Naumacos Polifemo 3D system (the evolution of L1) produces detailed 3D models, maps and plans of terrestrial and underwater archaeological areas. Ancient sculptures, pottery, underwater wrecks, cities, can be sampled before, during and after the excavation. The system, can work up to 300 meters depth it makes in motion 30 scans in a second merging them in just one 3D model.
The precision of the system is 0.05 mm and the resolution is unlimited, it can scan several squared meters at time. Contrariwise of the traditional terrestrial 3D scanner, the Naumacos Polifemo 3D can scan moving objects, it can also scan in motion itself, NP 3D System doesn’t need to be calibrated nor it needs to be placed any targets for its use. The operators are free to move while scanning the area of their interest. The result is a 3D model of all the surfaces sampled, and, at the same time a 3D motion capturing. By making a complete sampling of the site it will produce a tridimensional digital map of the superficial layer and a detailed 0.05 mm mapping of the position of the US. During the excavation the scan of each layer will record every single event and modification of the site and its condition. In the underwater environment it’s not possible to take pictures from several meters due to the light absorption and for the low visibility through the water, the same is by scanning a surface of underwater objects so it’s important to make several shootings from different angles. It is very important to line up each single scan accurately to avoid the possibility to add mistakes during this phase.

3. Conclusions

The project Restoring Underwater gave us the possibility to test the effects of different conservation and restoration methods on different materials in various environmental contexts, and has allowed us to assess and classify the effects of the different degradation agents and to understand, especially, the importance of ordinary maintenance.

This project demonstrates that the conservation and museum display in situ of underwater heritage must not just be considered an appropriate choice, but may in itself provide a strong stimulus for experimenting new materials and technologies, as well as representing an important factor in the socio-economic development of the communities involved.

In our opinion the three-dimensional laser scanning survey plays a prominent role in the planning and finished stage of the restoration in order to show the restored piece in a museum or even in the underwater park. The three-dimensional laser scanning of archaeological structures before and after restoration is basic in the identification and mapping of different typologies and stages of the decay. Indeed, the 3D surveys allow the management of the park to have a realistic vision of the situation of the archaeological site and to better quantify the financial commitments required to program the restoration works and maintenance of safety standard for visits. This also leads to a better use of resources that, as often happens, are particularly poor for the preservation of underwater cultural heritage.

To conclude, this method of 3D documentation shows better the state of conservation of the monuments, and increase the value of scientific dissemination.

The restoration campaign conducted by our Institute in September and October 2012 in the Villa con Ingresso a protiro has produced the execution of new survey using L1 in some rooms architecturally more complex to test and evaluate the possibilities of this 3D system.

References


of Photogrammetry, Remote Sensing and Spatial Information Sciences 34.5 (W12): 95–100.


Post-Excavation Analysis in Archaeology Using 3D-Technology: the Case Study of Hala Sultan Tekke

Kostas Anastasiades  
Vrije Universiteit Brussel, Belgium

Sorin Hermon, Nicola Amico and Giancarlo Iannone  
The Cyprus Institute, Cyprus

Karin Nys  
Vrije Universiteit Brussel, Belgium

Abstract:
Recent developments in survey techniques gave to the archaeologists new tools for documentation, analysis and visualisation of archaeological data. Laser Scanning, photogrammetry and topographical devices and their integration represent powerful tools to improve the understanding of archaeological sites and excavations. The archaeological documentation moved from a 2D representation (maps, sections, images etc.), in some cases not accurate, to an interactive 3D visualisation. In this paper we investigate how new technologies can optimise the post-excavation processing of past archaeological campaigns based on the case study of the Late Bronze Age harbour town near Hala Sultan Tekke (Cyprus). The focus lies on the correction of the original ground plan of the site which was drawn by hand in 1999 and never revised. Topographical surveys with DGPS and Total Station and Terrestrial Laser Scanner were executed to evaluate case by case the encountered problems and to develop an accurate record of the excavation.

Keywords:  
3D Modelling, Laser Scanning, DGPS, Integration Of Techniques, Ground Plan, Post-Excavation Analysis

1. Introduction

In recent years there has been considerable interest by archaeologists in three-dimensional documentation and visualisation. Total Station, GPS, Laser Scanner, photogrammetry are increasingly used as a method to register archaeological data and to analyse spatial information. These techniques offer new opportunities to investigate the documentation and can be used to revise and rectify old data (site plan, sections etc.).

In the case study of Hala Sultan Tekke we focused the attention on the correction of the original site plan, drawn by hand in 1999. We executed a Total Station and DGPS survey and did a digital acquisition of a part of the excavation area, representing an interesting study to understand the potential of interactive and immersive visualisation to analyse and interpret the data.

1.1 The Case Study: Hala Sultan Tekke

The Late Bronze Age harbour town near the mosque of Hala Sultan Tekke (HST) in Larnaca, Cyprus, was situated along the shore of a lagoon, the present day Larnaca Salt Lake, which served as its harbour. HST flourished particularly in the last part of the Cypriot Late Bronze Age (LC IIIC-LC IIIA, 13th-12th century B.C.). Several finds of pottery, bronze, copper and other goods indicate its important role in the Cypriot copper production, as well as many international trade links among which the most important ones lead to Egypt, Anatolia, the Aegean and the Levant (Karageorghis 2002, 110-112).
After two short expeditions by British teams at the end of the 19th century, Paul Åström (Gothenburg University, Sweden) started the scientifically sound archaeological investigations of HST in 1971. The last campaign was held in 2005. Most of the campaigns until 1979 have been published in the series *Studies in Mediterranean Archaeology* (Åström et al., *Hala Sultan Tekke* vols 1-12), but unfortunately, as from the 1980s, the excavation campaigns were not fully reported anymore.

The ground plan of the site (Fig. 1) was drawn by hand in 1999 and never revised. In this research paper we will try to assess the error in the hand drawn ground plan using data from a Total Station (TS) and Differential Global Positioning System (DGPS) survey and from a Terrestrial Laser Scanner (TLS) survey. To do this we will first shortly explain the data obtained from these surveys. The data from both surveys will subsequently be compared to the original ground plan, but also to each other in order to estimate and compare their own accuracies. Currently the research at HST (the site was named after the mosque) is focussed on the area of compounds C and H, which is marked in blue in Fig. 1. As such the site surveys were also focussed on this area of about 15 m wide and 30 m long.

Up to now archaeological research, or more generally heritage research, in which new technologies like TLS have been applied, are mostly focused on the acquisition of the data themselves. Having the model for mainly 3D visualisation purposes seems a very important issue (“3D-Scannen statt Zeichnen – Anwenderbericht.”, Wei et al., “3D Documentation and Preservation of Historical Monument Using Terrestrial Laser Scanning,” p. 73-90., Beraldin et al., “Establishing a Digital 3D Imaging Laboratory for Heritage Applications: First Trials.”, McGregor et al., “Two Avenues for Data: Rosslyn Chapel as a Terrestrial Scanning Case Study.”, etc.). This research paper is the first step within a broader research framework of using 3D models as information sources. With 3D scanning being the high-end technology for site-surveying we will investigate how we can extract the ground plan of the archaeological site near Hala Sultan Tekke from these models. Afterwards we intend to investigate the possibilities (and limits) of the 3D models in GIS applications.

Figure 1. Ground plan of Hala Sultan Tekke, drawn by hand during the Swedish Cyprus Expedition (Markou 1999), the area of focus is marked in blue.

Figure 2. Trimble 5600 (Trimble Navigation Limited, Trimble 5600 Total Station - Datasheet).
2. Field Work

2.1 Total Station and DGPS

In 2009 The Cyprus Institute performed a TS and DGPS survey at the site of HST. The TS which was used is a Trimble 5600 (Fig. 2). This device has the following main specifications (Trimble Navigation Limited, Trimble 5600 Total Station - Datasheet):

- DR (Direct Reflex) technology
- Range in DR-mode: up to 1600 m
- Accuracy: 2-5 mm

The Leica SR530 DGPS (Fig. 3) was used for the geo-referencing of the TS base point. The main specifications of this device are (Leica Geosystems AG, Leica SR530 Geodetic RTK Receiver.):

- Dual-frequency receiver
- On-board RTK (Real-Time Kinematic)

- Used as a real-time reference or for static observations at remote sites
- Compatible with Total Stations
- Accuracy: ± 25 cm

During the site survey at Hala Sultan Tekke (HST) only one base point was used for the TS. The TS base point was geo-referenced with the DGPS, and this point was also chosen for the base GPS station. With the rover receiver four other GPS points were taken so all the TS measurements could also be geo-referenced with an accuracy of 2 cm. Fig. 4 shows the measured points superimposed on the hand drawn Swedish ground plan. The purple points are the GPS points measured with the rover receiver. The yellow points indicate the pickets which were measured. These are very accurate points in this sense that we know exactly where they are on the site. They can also be pointed out on the ground plan, because they represent the grid. The black points are points on the

Figure 3. Leica SR530 (Leica Geosystems AG, Leica SR530 Geodetic RTK Receiver).

Figure 4. Total Station survey, performed by STARC, shown with the aligned Swedish hand drawn ground plan as underlay. Yellow: pickets, black: points on the walls, purple: DGPS points, red: base GPS and Total Station base point, green: levels.
walls. They are much more difficult to define on the ground plan, because we are mainly dealing with rubble walls. For this reason it is also hard to retrace them on the site.

Now that we were able to align the hand drawn Swedish ground plan with the TS and DGPS measurements, the discrepancy between the two becomes clear. We can measure an error in the hand drawn ground plan of up to 75 cm, which is enormous. With this information the need for a new, accurate ground plan has been assessed and demonstrated.

2.2 Terrestrial Laser Scanner

Two TLS surveys were executed by The Cyprus Institute on the site of HST. The first one in 2009, the second one in 2011. The 3D scanner with which the first survey, in 2009, was performed was a Z+F Imager® 5003. This laser scanner was designed in 2002 and has the following main specifications (Zoller + Fröhlich GmbH, Technical Data Z+F Imager® 5003):

- Phase Shift scanner
- Scans up to 500,000 points per second
- View range: horizontal: 360° vertical: 310°
- Accuracy: ± 3 mm

In the second survey, in 2011, a Surphaser 25HSX was used. This 3D scanner was built in 2010. It’s main specifications are listed below (Basis Software, Inc., Surphaser - 3D Laser Scanners and OEM Products for Laser Scanner Manufacturers):

- Phase Shift hemispherical scanner
- Scans up to 1.2 million points per second
- View range: horizontal: 360° vertical: 270°
- Accuracy: ± 0.5 mm
One cannot help but remark the improvements the technology has undergone in less than a decade.

After the on-site survey all the acquired point clouds were aligned in order to gain the complete site-model. Subsequently the obtained 3D model was geo-referenced by aligning it with the TS points. By doing this the relative coordinates of the point clouds were transformed into absolute coordinates, orienting the model to North and allowing to retrieve the levels of different parts of the site, just by reading the absolute coordinates of the relevant points.

From this point on the 3D model can be used for different purposes. All the remains of the Hathor temple in Naga, for example, were scanned from all sides to perform a virtual 3D reconstruction, placing the different fallen and broken down stones back in place (TRIGONART, BAUER PRAUS GBR 2010). Accurate 3D measurements can be obtained from the 3D model, as well as level measurements when the 3D model has been geo-referenced. In turn these measurements can help to accurately map the damage historical monuments/heritage sites have suffered over time, which is important for consistent restoration-planning (Novello, and Marchis, “3D Models for the Restoration Project: Some Issues and a Case Study.”, Wei et al., “3D Documentation and Preservation of Historical Monument Using Terrestrial Laser Scanning”, etc.).

3D scanning will indeed be of great value for these purposes, but in essence the models are then only used for viewing and for measuring. The measured lengths are the only real information extracted from the models.

However, an interesting aspect about these 3D models is that one can make both horizontal and vertical sections through them. The horizontal sections can be made at any height, the vertical sections can be made anywhere along the length of the walls. In Fig. 9 the Swedish ground plan is shown and aligned on it the TS and DGPS survey (blue) as well as a horizontal section from the 3D model (red). We can see that the TS measurements and the section from the 3D model are perfectly aligned, which not only emphasises again the errors in the Swedish hand drawn ground plan, but they are also a validation of each other: two distinct devices generating perfectly matching results and as such proving the accuracy of their measurements.

As interesting as this observation may be, the section of the 3D model is not very useful as a ground plan, because it does not show the build-up of the walls. Another way of looking at the 3D model is in orthoview mode. The top view of the 3D model generates in this way something like a ground plan (Figure 10. Orthoview of the top of the 3D model). Also the elevations of the walls can be visualised in this way. One
must keep in mind, though, that this orthoview is a raster image. For the elevations of the walls this may be sufficient, but for a ground plan a vector file is more convenient, since a vector file can easily be modified in CAD or GIS software, for example to show the different building phases of the site (since we are dealing with a horizontal stratigraphy). Also the orthoview is like an aerial photograph of the site, providing too much detail with regards to the ground plan (including all the dirt like weeds which were still present on the site during the scanning sessions). Our investigation of the opportunities and limits of the 3D models will as such continue with looking for an automated or semi-automated way to extract a real vectorised ground plan from them.

3. Conclusions

In this research paper we have assessed the accuracy of the hand drawn ground plan of the Late Bronze Age site of Hala Sultan Tekke. The tools we had to do this were the data obtained from a TS and DGPS survey, as well as a TLS survey which were executed on the site by STARC.

First we discussed the TS and DGPS survey in terms of the data we obtained with it. These data are the points which were measured during the survey. Combining the TS and DGPS measurements results in a geo-referenced grid in which other survey data can be aligned, in order to geo-reference those as well. The hand drawn ground plan of HST was then aligned with the georeferenced TS measurements, which revealed an error in the plan of about 75 cm. With this information the need of a new, accurate ground plan was established and demonstrated.

Subsequently, we discussed the TLS survey. The 3D pointcloud from TLS was aligned within the grid obtained from the TS and DGPS dataset in order to geo-reference it. Very interesting is the ability to make sections through the 3D model. We made a horizontal section and aligned it with the hand drawn ground plan, as well as the TS and DGPS measurements. Doing this emphasised again the error in the hand drawn ground plan, but more interestingly, the section from the TLS model was perfectly aligned with the TS and DGPS measurements. This proved both their accuracies, since they are two distinct devices generating perfectly matching results.

TLS surveys are very elaborate site surveys. The 3D model gained from the 3D scanning is in fact a full 3D representation of the archaeological site. After all, practically every possible point has been measured by the 3D scanner in order to create the point clouds. Any measurement can be done perfectly and more accurately in the 3D model, because it is in real-life scale. The continuation of this research will thus be to find an automated or semi-automated way to generate the ground plan from this 3D model.

References

Post-Excavation Analysis in Archaeology Using 3D-Technology: the Case Study of Hala Sultan Tekke
Kostas Anastasiades et al.

Architectures, Trento, Italy, March 2-4, 2011.


A New Approach for Interactive Procedural Modelling in Cultural Heritage

René Zmugg, Ulrich Krispel, Wolfgang Thaller, Sven Havemann, Martin Pszeida and Dieter W. Fellner
Graz University of Technology, Austria

Abstract:
We present a novel approach for the efficient interactive creation of procedural 3D models for creating synthetic 3D reconstructions in cultural heritage. The benefit for the CH community is a 3D modelling tool that scales better than conventional forward modellers (like SketchUp) but does not require actual coding (like CityEngine). Our tool uses the split grammar approach, but it allows for code-free, direct model manipulation in 3D. To build up the procedural model we propose a modelling by example method that combines the intuitiveness of direct 3D interaction with almost the flexibility of programming. The user can interactively browse through the refinement hierarchy, apply rules and change parameters visually. The tool provides only a small set of well-defined modelling operations. The question is: Are they sufficient for the domain of classical architecture? For our case study we have chosen a prime example of classical architecture, the Louvre.

Keywords:
Procedural Modelling, Reconstruction, User Interface

1. Introduction

There is no such thing as a perfect 3D reconstruction of a historic building. Any 3D model of a real building is necessarily an abstraction and incorporates a certain degree of interpretation (depending on the purpose of the model). A laser scan of a building includes a lot of detail but very little interpretation. For example, small flaws on a surface will be recorded accurately, but the idealized flat surface of a wall will not be represented at all. In a polygonal model, surfaces are interpreted as flat polygons; details can be visualized as textures or be discarded if they are considered irrelevant. Procedural models constitute the extreme end of this scale. We attempt to “explain” a buildings shape by a set of rules specifying amongst other things regularities or symmetries; details that do not fit this structure can be represented as an exception to these rules, or be neglected in return for modelling efficiency.

Compared to manual polygonal modelling, procedural modelling has the potential to:
• increase modelling efficiency by exploiting regularities
• allow for post-facto parameter changes
• increase understanding of a model through experimentation with different interpretations
• yield a hierarchical interpretation of a model that can be progressively refined

Procedural interpretations of a shape can be arbitrarily complex, which is why full-featured programming languages have often been used to create procedural models. We have developed a tool that consists of a set of basic operations which the user can interactively apply to a 3D model; the procedural description is created in the background without requiring the user to do any programming. By focusing on specific structures predominant in architecture (straightforward repetition and symmetry) the common cases can thus be handled in a simple and efficient manner. This does not limit the expressiveness of the system because writing program code is still available as a fall-back solution for more complex details.
As an experiment, we have decided to apply our system to reconstructing parts of the Louvre, as it is a prime example of classical architecture. Additionally, it is of non-trivial size and complexity and allows us to evaluate the scalability of our approach.

1.1 Contribution and Benefits

This paper has two novel technical contributions, the parametric assets (Section 3.2) and direct parametric modelling. On an abstract level, the challenge is to find a bridge between the flexibility of programming and the intuitiveness of direct 3D interaction. We propose a “modelling by example” method, providing a small but well-defined set of modelling operations.

The benefit for the community is a more efficient 3D modelling toolkit than classical forward modellers (SketchUp (Google 2012), etc.) without the overhead of coding (CityEngine (Esri 2012)).

2. Related Work

Procedural modelling is the umbrella term for methods to describe a 3D object using a set of rules or an (often domain specific) programming language. Many methods are based on formal language theory, which studies formal grammars and languages. Introduced in 1968, Lindenmayer systems, or L-systems (Prusinkiewicz and Lindenmayer 1990), were an early representative of this methodology. A set of string rewriting rules were used to model plant growth processes.

Shape grammars have since then emerged as a formal method to describe the structure of complex man-made shapes. Stiny and Gips pioneered shape grammars as a formal method not only for 2D shapes (Stiny and Gips 1972), but also for 1D and 3D (see: (Stiny 1976) and (Stiny 1980)). Shape Grammars, in the sense of Stiny, can be additive, subtractive, or subdividive. The very first grammar by Stiny was based on subdivisions (Stiny 1977).

Recently, their method was developed further for the modelling of 3D architecture by Wonka et al. (Wonka, et al. 2003) who introduced split grammars. It was further improved by Müller et al. (Müller, et al. 2006) up to automatic creation of whole cities, which was commercialised as CityEngine by ESRI (ESRI 2012). Editing a textual description of the rule set, however, is not the most intuitive way to describe a building. An interactive visual grammar editing method that allows interactive local rule modification was proposed by Lipp et al. (Lipp, Wonka and Wimmer 2008). They also developed methods for semantic and geometric selection, and for describing local rule exceptions. Our system achieves a similar expressiveness but uses a less restricted modelling method, seemingly similar to conventional forward modelling. Recently, Patow (Patow 2010) introduced a change in the paradigm for editing shape grammars from a set of rules to a graph-based representation, which is more intuitive. Still, the modifications are carried out on the meta description (the graph) instead of the concrete 3D model.

Methods that build upon general purpose languages are less strict and more
general than formal methods. Grasshopper 3D is a visual programming language for the Rhinoceros CAD system (McNeel 2012). It allows building generative models in a dataflow fashion. Furthermore, every larger 3D modelling package is equipped with a scripting language that is able to procedurally generate 3D content. Examples are C# for Revit, AutoLisp for AutoCAD, MEL for Maya (all Autodesk (Autodesk, Autodesk 2012)), Ruby for SketchUp (Google 2012) etc. Our goal, however, is to obtain procedural expressiveness by direct 3D model manipulation. We use the generative modelling language (GML), a procedural modelling language that was inspired by the PostScript language from Adobe, but incorporates operators for 3D shape design instead of typesetting operators (Havemann 2005). Our GML code, however, is generated in the background and not by manual coding, but instead by a dataflow graph based representation for procedural models (Thaller, et al. 2012).

Recently, Chevrier et al. (Chevrier, et al. 2010) examined the practicability of parametric components in heritage reconstruction. Their components are supplied by a plugin for the Autodesk Maya environment. They are instantiated using an initial estimate provided by a user and can be further refined with the help of photographs, plans and point clouds. Our workflow is inspired by this excellent targeted system, but attempts at being more general.

Our reconstruction is not the first attempt at a procedural reconstruction of the Louvre. Calogero et al. generated two alternative proposals of the east wing of the Louvre (Calogero and Arnold 2011) using procedural modelling (CityEngine). Their split analysis of the Le Vau/Perrault/Le Brun design of 1668 of the east wing as it stands today served as basis for our reconstruction.

3. Parametric Modelling System

A parametric model is a model whose parameters can be changed after the modelling process is complete (“post facto”). For a given set of parameter values, a concrete three-dimensional model instance can be automatically generated from the parametric model. The parametric model is a recipe for creating a concrete model from a given set of parameter values.

In our system, modelling is done “by example”. The user always sees a concrete model on the screen, and the system automatically keeps track of the operations used to generate it. When a parameter is changed, the affected operations are automatically re-evaluated; the concrete model is updated.

Our parametric modelling system can be understood as being made up of three layers. At the lowest layer, the system manages volumetric elements and provides modelling operations to edit these volumes (sections 3.1 and 3.2). The middle layer comprises automatic GML code generation. The top layer is the graphical user interface that allows a user to combine the modelling operations to build parametric models (section 3.3).

3.1 Basic Modelling Operations

Modelling in our system typically proceeds in a top-down manner, starting off with a coarse approximation and subsequently refining the result. The basic modelling entity is a volumetric element called the scope; every scope has an attached local coordinate system (LCS), which the modelling operations operate on. The basic operations are “split” to partition a scope into a certain number of sub-scopes with given lengths, and “repeat”, to partition it into as many scopes of the same given length as possible.

Our system provides a set of twelve basic modelling operations (not counting the
Figure 2. A scope (a) is split into four parts along a given direction (b). The size of each part can be defined using relative proportions or absolute measurements. The (red) discs can be dragged to change the split intervals. Picture (c) shows the application in the reconstruction process. This interval split was used to split the façade into horizontal segments in an early partition step.

Figure 3. Extrude. The selected scope (a) is extruded in the given direction producing a new scope (b). Dragging the (blue) discs modifies the extrusion offset interactively. Image (c) shows the application in our reconstruction context. The extrusion operation was used together with the diagonal split operation to model the ledges.

Figure 4. Repeat. The repeat operation splits a scope (a) into equally sized parts, given a minimum size and a direction (b). The space is distributed evenly among the elements. Changing the minimum size can result in a change of the number of elements. The resulting elements are all contained in the same link group; therefore modifying any of these will result in modifying all. Dragging the middle (red) disc modifies the minimum size of the result scopes. The sequence of pillars was realized using this operation (c).

Figure 5. Link allows to apply operations to groups of objects: Two scopes have been linked (a), and an interval split operation is applied to the selected scope. The linked scope automatically mimics the same split; three pairs of linked scopes are generated from one linked pair (b). The decoration above the windows was realized using link (c).
parametric assets described in section 3.2). We have separated these twelve operations into four groups: creation operations (generate new scopes), refinement operations (create sub scopes), grouping operations (treat scopes equally) and modification operations (change scope state).

In classical architecture, repetition is common. Therefore, it is often necessary to apply the exact same operation to several scopes. Therefore, our system supports linking multiple scopes together (link group) to reduce the workload. When an operation is applied to one scope of a link group, it is automatically applied to all other scopes as well. These operations are all applied using the scope’s respective LCS; preceding modifying operations (e.g. mirror) on a single scope of the link group still affect only that single scope. For example, mirroring the LCS of a scope and linking it to another scope can be used to model symmetry. The difference to conventional modelling systems is that the geometry is not simply copied and scaled, but evaluated in the current scope’s state, e.g. a parametric column is applied to a link group of scopes with different heights.

In the remaining part of this section we will explain our modelling operations in detail. Each example consists of a schematic, abstract illustration as well as its application in a concrete example. The concrete examples are taken from our reconstruction of the Louvre east wing pavilion. Directions are defined in terms of the LCS of the selected scope and are in most cases the X, Y or Z axis directions.

The creation group contains operations for scope creation and deletion:

**Box.** Create a box-shaped scope: A box-shaped scope (of adjustable dimensions) is created as a starting point for the modelling process.

**Void.** Make a scope invisible: The opposite of Box. The selected scope is rendered invisible and cannot be refined further (but it still exists internally).

The refinement group provides operations to partition a scope in smaller scopes and to add detail.

**Interval Split.** Split a scope into parts with a given list of sizes: The selected scope is split along a given direction into smaller parts. Planes orthogonal to the given direction are used for splitting. The sizes of the subparts are defined by a set of interval lengths along the split direction. These can be defined as combinations of relative proportions and absolute measurements. Figure 2 illustrates the usage of this operation.

**Extrude.** Displace faces of a scope in a given direction: The faces of the selected scope with a normal similar to the given direction are displaced by a certain given amount. A new scope is created for the extruded part. Figure 3 shows the application of the extrude operation.

**Diagonal Split.** Split a scope along one of the diagonals of the scope’s bounding box:
Introduces a diagonal split to the selected scope and divides it in two parts. There are six possible cases; the selected case is specified as parameter. Figure 3(c) shows applied diagonal splits in the extrusions.

**Repeat.** *Split into linked parts of given size:* The selected scope is split into the maximum number of parts along the provided split direction with the given minimal size. The resulting scopes form a link group (Fig. 4).

**RepeatABA.** *Split into alternating sequence of parts A and B of given sizes:* The selected scope is split along the provided split direction into an alternating sequence of parts with the given two sizes for parts of group A and B. Either size of the parts of the group A or B can be set to a fixed size, any remaining space is distributed evenly among the elements of the opposite group. The sequence begins and ends with an element of the group A. The results are two link groups; all A parts are treated the same, and so are the B parts. For example, automatic distribution of pillars and gaps in a scope.

**Merge.** *Merge scopes into one:* Several selected scopes are merged into one scope. This is the inverse to the refinement operations.

Thirdly, the grouping operations provide methods for processing several scopes at once.

**Link.** *Treat these scopes in a similar way:* A link group is created from the selected scopes. Any operation applied to one scope of the group is automatically applied to the other scopes as well in their respective LCS. Figure 5 shows the behaviour of linked scopes.

**Specialize.** *One (or more) scopes in a link group are exceptional:* This is the inverse of the link operation. One or more scopes are released from the link group. Formally, the link group is partitioned into two disjoint link groups. Each group can then be refined in a different way.

The following modifying operations do not change the scope geometry, but only in its state, e.g. its LCS. All subsequent operations on the scope are influenced by these changes.

**Mirror.** *Mirror the orientation inside a scope:* Modifies the LCS, given a mirror direction (X, Y or Z axis of the LCS). Figure 6 shows the behaviour of mirrored scopes.

**Rotate.** *Rotate the orientation inside a scope:* This rotates the LCS by 90 degrees given a rotation axis (counter clockwise or clockwise).

It is important to emphasize the difference between the merge and link operation. Figure 7 illustrates this difference. Linked scopes behave similarly, so if a linked scope is split in half, so is each scope in the link group individually. The splitting plane may differ when the scopes have different sizes. Merged scopes have become one single scope; therefore, there is only one splitting plane.

Additionally, the system offers a method to give names to specific values that can be
referenced by operations. Using these labelled values it can be ensured that several operations use the same input value (e.g. an extrusion depth).

3.2 Parametric Assets

In addition to the low-level modelling operations described above, our system includes a set of pre-modelled assets, i.e. a set of ready-made windows, arches, and other architectural elements that can be inserted into a façade (Fig. 8). They are, however, not the usual static 3D models but parametric models. Thus, they can approximate a wide range of actual architectural elements by adjusting various parameters and adapt automatically to different sizes and proportions when applied to differently shaped scopes; so one should think of them as operations, rather than of objects.

The asset library is extensible; our parametric assets are actually small GML programs. While GML is harder to use (it requires coding), it is also more powerful than our interactive system. Therefore, anything too complex to be conveniently modelled using the GUI can be scripted as an asset beforehand.

The distinction between low-level operations and high-level parametric assets is to some degree arbitrary; some asset operations can also be seen as more sophisticated split operations, as they partition a scope into parts that can be refined later on. One example is the round-hole operation: With a given minimum outer distance, the operation partitions the scope into a cylindrical scope (whose radius is calculated automatically) and the remaining outer part as a second scope. Such parametric partitioning assets can be very versatile. The round-hole operation can be applied subsequently to produce detailed architectural elements such as windows with borders. It can also be used to produce a round pillar structure when applied to a rotated scope (see Figure 9 for reference).

3.3 User Interface

The user interface of our system (Fig. 10) consists of two main parts, a 3D view (A) that shows a concrete 3D model, and a side bar (B) with a “toolbox” of operations. Scopes can be selected with the mouse. Intermediate results, i.e. scopes that have already been replaced by other scopes during the refinement process, are

![Figure 8. Various parametric assets from our asset library.](image)

![Figure 9. Versatile usage of the round-hole operation: the same operation can be used to create a round window (left) by voiding the inner part and to create a pillar (right) by voiding the outer part.](image)
not visible as such in the 3D view, but they can still be selected. This is done by first selecting one of the more refined scopes and then turning the mouse wheel to navigate in that particular scope’s hierarchical refinement history.

To illustrate this, in order to select the scope that represents the entire first floor of a façade, one might first click on a window pane of a window in the first floor and then turn the mouse wheel to get first to the scope that encloses the entire window, and then to the entire floor. Turning the wheel further will finally select the entire façade. In cases where a scope has more than one ancestor (e.g., when it was created using merge or link), the ancestor is chosen to which the mouse pointer points.

Operations can be applied to the selected scope using the buttons in the side bar. While a scope is selected, the parameters of the operation shown in the respective property box can still be changed. Depending on the operation, a three-dimensional widget (a “manipulator”) is shown in the 3D view to allow interactive parameters adjustment. For example, if the user selects a scope that is the result of splitting a larger scope in two parts, a bar-shaped manipulator is shown with a sliding disk for the user to change the split line (Fig. 10).

4. Reconstruction

4.1 Source Material

When reconstructing a building, the first task is to acquire sufficiently accurate source material. If not enough measurements exist, a reconstruction with conventional 3D modelling software is problematic. We have found that using a conventional modelling tool, such as Google SketchUp (Google 2012) correcting inaccurate measurements in the modelled geometry later on requires excessive time and effort. With a parametric model, all parameters can still be set to more accurate values at any time they become known.

In case of the Louvre we were provided with some of the original building plans, providing ground layouts as well as façade layouts for parts of the Louvre. In total we had access to eight ground plans and nine façade plans of different parts of the Louvre.
They cover parts of the Cour Napoléon and the Cour Carrée. One plan contained detailed information on profiles for the respective façade. Although the façade plans provide much detail, they lack the extrusion depths of the façade elements. The extrusion depths are therefore only estimated from photographs. Since incomplete information is typical in such reconstruction tasks, using a parametric approach is quite important.

4.2 Stepwise Reconstruction of a Louvre Façade

This section describes the methodology for creating a procedural reconstruction of the pavilion part of the east wing of the Louvre (Fig. 11). We follow the split proposal given by Calogero et al., compare Figure 5 in (Calogero and Arnold 2011).

At first, a rectangular block representing the pavilion part is partitioned into the vertical structure of podium, pilaster, entablature and parapet (Fig. 12(a)) using the interval split operation (Fig. 2). Furthermore, the parts which are not entablature or ledges are split into three horizontal parts, which reflect the coarse structure of the façade (Fig. 12(b)). Next, the parts on the right side are mirrored and linked to the left parts to reflect the symmetry of the façade (Fig. 6). Consequently, only one side...
Figure 13. Ledges are modeled using extrusion, diagonal splits and voiding the unnecessary parts. Note that the corner part is split twice using diagonal split (left column, top to bottom). This method can be used to model more complex profiles (right).

Figure 14. Modelling a pillar which is adjoined to walls with different depth offsets. First, a round-hole operation is applied (left). Then, the outer part is split into two parts (middle left), to which the different depth offsets are applied using another split operation (middle right). Finally, the outer parts are voided (right).

Figure 15. Rendering of the final step of the reconstruction. Special parts such as the round pillars, the parapet and linked ledges are highlighted.

A New Approach for Interactive Procedural Modelling in Cultural Heritage
René Zmugg et al.

has to be modelled. More detail is added to the façade parts by splitting down to the hierarchy level of windows and pillars (Fig. 12(c)). The correlation of the partitions of pilastrade and parapet was realized with interval splits using labelled values (see end of section 3.1).

We continued adding detail to the podium part first. The lower podium part is split off and extruded. The upper part is split into window and wall parts. The window parts are linked, due to the similarity of the three podium windows. Window detail is added by applying
a parametric window asset split and extruding the frame. The keystone is created by splitting the respective parts of the window frame and space above the window into three horizontally arranged parts, merging the middle part to form the keystone and extruding it. The window sill is created by splitting the lower window frame and extruding it sideways; the inner window part is created using a parametric crossbar asset (Fig. 12 (d)).

The ledge is modelled using a series of extrusions. First, the vertical partition is split and extruded to the front and sideways. The corner parts are modelled by again extruding the sideways parts to the front. The slanted parts are created using diagonal splits; corner parts are split twice using diagonal split (Fig. 13). The result can be seen in Figure 12 (e).

For the pilastrade part, the pillars are first split vertically into base, column and capital parts. The base detail is created using diagonal splits, and the fine alternating structure of the rectangular pillars is created using repeatABA. The window is created using parametric assets for the border and the top parts (similar to the bottom windows), see Figure 12 (f). The ledge above the window is created by linking the space between the pillars and above the window and applying an extrusion and splits (see also Fig. 5). The middle part is first split into depth direction, and the front part is voided to create the depth offset.

The window is modelled using parametric assets. Additional details are added using splits and extrusions. The pillars need special handling due to different wall depths on the left and right side of the pillar. We get the right geometry by splitting the space around the pillar into four parts, setting the corresponding depths, and voiding the outer parts as illustrated in Figure 14. The intermediate result can be seen in Figure 12 (g).

The capital part of the pillars is roughly approximated using parametric assets (round-hole and interval splits). The entablature is modelled similarly to the ledge above the podium, just with more extrusions. The solid parts of the parapet are linked, and their top part is extruded to create the ledge; the pillar rows are created using repeat (see also Fig. 4). The final result is shown in Figure 12(h).

This concludes the reconstruction of the pavilion part of the east wing façade of the Louvre. The achieved level of detail can be seen in Figure 15. Decorations (e.g. statues or ornaments) have not been modelled.

5. Reconstruction Results

This section provides information about the current status of the reconstruction. We concentrated on the following façades (highlighted in Figure 16):

- façades of the Cour Carée
- arcade section of the Cour Napoléon
- the outer façade of the east wing

To assess the practical applicability of our system we let a computer science student attempt a reconstruction of certain façades of the Louvre. The student had not been involved in the development of the system, so he was able to approach the task from an end-user’s perspective, except that the prospective end-
user will have a background in art history rather than computer science.

On the whole it is to say that about 2/3 of the reconstruction time is spent modelling fine façade details (extrusions and profiles), which suggests that there is much potential for improvement. Furthermore, (subjectively) much time went into testing out pros and cons of different workflows. The resulting speedup will have to be verified, and quantified, in the upcoming user studies. While the metric ground truth would be, e.g., a 3D Laser scan of the façade, a semantic ground truth is more difficult to obtain: Does our model contain all “relevant” architectural elements of the façade?

5.1 Façades of the Cour Carrée

Based on the available source material (see Section 4.1) roughly 80% of the façades of Cour Carrée could be reconstructed so far. The obtained level of detail is shown in Figure 17.

Naturally, the speed of the reconstruction depends on the degree of acquaintance with the modelling system. After some time to get used to the system, the student was able to reconstruct the façade of the Cour Carrée in roughly 10-12 hours. A coarse approximation was achieved faster, in a few hours, and the remaining time was spent modelling finer details. A rendering of the façade is seen in Figure 18.
5.2 Façade with Arcades at the Cour Napoléon

As for the Cour Napoléon, about 40% of the façades are finished. For the remaining parts a coarse façade layout is reconstructed. The arcade section of the Cour Napoléon was done in about 7-8 hours. To judge the obtained level of detail a comparison between photograph and model is shown in Figure 19. These images show that still some work is needed to achieve the same level of detail.

5.3 Outer Façade of the East Wing of the Louvre

The reconstruction of the pavilion part was discussed in detail in Section 4. It took us about 8 hours to reconstruct the pavilion. The whole façade containing the pavilion was reconstructed by the student mentioned above. He needed about 12 hours for the whole façade that is a little bit less detailed than our pavilion. Figure 1 shows the acquired result.

6. Conclusions/Outlook

We have presented a parametric modelling system that allows non-expert users to create more sustainable procedural 3D reconstructions of historic buildings. Unlike the Louvre use case, we anticipate the system to be most useful for speculating about destroyed and non-existing buildings: New archaeological findings might require adjusting parameters of a 3D hypothesis, which can be carried out much faster with a procedural model than with Google SketchUp etc.

Much remains to be done because this system is still in its prototype phase. For example, we experienced high memory usage when combining several detailed reconstructions of façade parts into one model.

So far, the focus in the development of our system has been on a compact set of basic operations and on means to combine them. For example, the details shown in Figure 19 (e.g. profile extrusions) have been modelled mainly using several extrude operations. While these operations are expressive enough to achieve the results, the amount of time spent modelling these details suggests that developing a special purpose tool for profile extrusions is practical. Several other operations are still missing or in development, these are amongst others adding roofs and incorporating 3D-scanned material (e.g. statues).

The Louvre reconstruction exercise has taught us that (a) the Louvre, especially its interior, poses enough interesting challenges to keep us busy presumably for several years, and that (b) although our approach is quite useful for buildings, we should systematically
A New Approach for Interactive Procedural Modelling in Cultural Heritage
René Zmugg et al.

examine its usefulness for other domains of man-made shapes.

Acknowledgements

We gratefully acknowledge the generous support from the European Commission for the research project 3DCOFORM (3D COllection FORMation, 3D-coform.eu) under grant number FP7 ICT 231809. We thank the Louvre for providing the source material for the reconstructions by facilitating access to their archives. Furthermore, we thank Erica Calogero for permission to use the photograph of the outer east wing (Fig. 11).

References


“Instant Architecture.” ACM Transactions on Graphics
Virtual Reality Simulations in Cultural Heritage

Ioanneta Vergi
Sainsbury Institute for the Study of Japanese Arts and Cultures, UK

Abstract:
While virtual technology has evolved into a popular and user-friendly medium for people to experience heritage, the shape of tangible heritage has progressively been converted into a digital reality. The role that museums and cultural heritage institutions play in interpretation and representation using new and emerging virtual media is strongly involved with the adoption of simulacra, digital pictures and signs. Virtual heritage reconstructions have come to stand as displays in their own right, almost replacing the historical role of objects in exhibition making. In a sense they have become objects and heritage is seen to be important as it is more accessible to the public when it contains and gives information that is communicated by a variety of new technologies such as VR. Example for the visualization of historical knowledge thought virtual reality will be taken from the Virtual Kyoto project created by the Ritsumeikan University, Japan which aims to create city models where visitors can travel virtually in the historical city of Kyoto.

Keywords:
Virtual Heritage, Simulations, Virtual Kyoto, Public Engagement

1. Introduction

The use of virtual reality (hereafter VR) systems in representations of cultural heritage sites has evolved into a powerfully developing application2 for three dimensional (hereafter 3D) graphics and heritage reconstructions. VR in heritage applications is an electronic-simulated space that uses 3D computer models to duplicate or reconstruct historic environments. Cultural heritage with the use of VR can be defined as the reconstructive recording of cultural properties within virtual technologies and is usually referred to as ‘virtual heritage’ (Tan and Rahaman 2009, 144).

Although this practice is used to depict the way in which monuments, group of buildings or artifacts could have been built, it runs the risk of being accepted as reality and persuading the audience to accept a particular visualization. Even if the viewers are aware that the depicted site cannot be found in the same condition or does not exist at all, the power of the VR can still deceive viewers’ minds and establish specific pictures in their memories. Hence, the viewer can be led to a certainty which never existed in reality.

Virtual reality that uses special graphics, video images and stereo sound to enable people to interact with the world in different ways via a computer, has been made widely popular through media and video games on a large scale. Its impressive, wide-ranging, almost infinite uses, means that it has also been significantly taken up to promote diverse businesses from designing planes to training employees. However to depict the process of heritage reconstructions through virtual reality is problematic. With a view to understanding the past as a cultural reconstruction in the present, this paper follows a theoretical approach to discuss and explore a virtual construction strategy within the specific cultural context of Japan and the creation of Virtual Kyoto by the Ritsumeikan University.

While virtual technology has evolved into a popular and user-friendly medium for people to experience heritage, the shape of tangible heritage has progressively been converted...
into a digital reality. Through virtual heritage reconstructions, the approach to heritage is strongly involved with the absorption of simulacra\(^3\), digital pictures and signs (Flynn 2007, 349). The conversion of heritage to digital systems means that virtual interpretation has been embodied in the value and experience of heritage.

It is the intention of this paper to argue that virtual heritage could not be recognised as a reconstructed simulation of the past, but rather as a comprehension of heritage in the present. Moreover, the papers illustrates how historical based VR environments such as the Virtual Kyoto project could support the cultural heritage purposes of historical teaching and learning.

2. Reconstructing Virtually the Past

Reconstructions of heritage sites in virtual environments are produced by cultural organisations, museums and universities and usually are presented through the use of Internet and interactive exhibitions (Zara 2004, 101). VR spaces in cultural heritage applications are different firstly from the online VR worlds of which a multitude can be found on the Web and secondly from the profit-making video games with historical themes.

A good example of online virtual worlds is Second Life developed by Linden Labs with a count of thirteen millions users worldwide (Anderson et al. 2010, 257). It is the largest online platform user-created; the biggest 3D virtual world community that simultaneously represents virtualized real cities and heritage sites with imaginary words. However the extended use of fantasy time and spaces cannot allow Second Life, and similar virtual reconstructed worlds, to be part of the ‘serious’ applications for cultural heritage.

Since the late twentieth century, the utilisation of immersive virtual platforms\(^4\) in heritage interpretation in museums and cultural organizations has developed to an advanced state of scientific interplay which has inevitably affected the traditional obsolete display of monuments and artifacts. It is worth noting that what is being promoted most of the time is not heritage sites that can be found in the present but environments that have been destroyed or are not available in their original form.

In terms of the relation between tangible and virtual heritage, the VR replicas could evolve into a more acceptable and persuasive form of heritage than the material heritage itself, as the presentations of virtualised reconstructions make the digital models more visually appealing to a wider public (Barcelo, 2002). However, the new ways of experiencing the past can give to the viewer the impression that virtual projects literally exist in specific representations.

Forte and Siliotti (1997) in Virtual Archaeology put forward the ability of humans to be embodied in a virtually 3D heritage space and therefore the mind-body capacity to explore cultural heritage through a different degree of experience. But if spectators are able to explore cultural heritage and the past in the cyber space of virtual heritage to what extent can they experience it?

Lowenthal (1985) in The Past is a Foreign Country examines how the ever-evolving and changing past shapes the life of contemporary societies. He states that cultural heritage sites and artefacts cannot be reconstructed in the present and therefore monuments or sites when they lose their very first form cannot re-establish their resources. According to his concept, the past is an accomplished fact and as a result reconstructed sites cannot be perceived as authentic representations of the past.

---

\(^3\) According to Postmodernist French social theorist Jean Baudrillard a simulacrum (plural: simulacra) is not a copy of the real, but becomes truth in its own right, the hyperreal.

\(^4\) Immersive virtual platform is the environment where spectators disconnect from the exterior physical presence and dive in a total artificial space, usually with the use of stereoscopic glasses.
However, even the most accurate and studied heritage reconstruction could never become equal presentations of the prototypes. Benjamin (1969) in his essay ‘The Work of Art in the Age of Mechanical Reproduction’, constitutes the reflection by way of which the interpretation and experience of culture is changed by the consumption of what was considered at his time, modern technology. He focuses on the media of photography, film and advanced printing techniques and argues that the mechanical and automated methods that these new technologies use to reproduce and re-present a specific part of culture or artwork, have destroyed its ‘aura’. His concept of aura characterises of the existence of the work of art in a unique place and time and therefore he emphasises that “even the most perfect reproduction of a work of art is lacking in one element its presence in time and space, its unique existence at the place where it happens to be”.

Even if Benjamin was analysing this concern before VR had been discovered, taking a work of art as his centre of attraction instead of the general notion of cultural heritage, the relationship he demonstrates between advanced forms of technology with the concept of re-making the sense of place and time is a process whose meaning and significance reflects effortlessly on virtual reality. Thus, virtually reconstructed spaces cannot be then representative replicas because they do not have the physical site to duplicate this accuracy.

Enhancing this argument, Lowenthal (1985) states that ‘the more realistic a reconstruction of the past seems, the more it is a part of the present’. Therefore, the documentation and presentation of heritage sites in virtual simulations cannot be recognised as original images but as comprehensions of their past in the present. Nevertheless, there is always the possibility to reproduce virtual environments by merely depicting what already exists. However, the purpose of VR in the service of cultural heritage is not to depict what is there but what is unable to be seen, otherwise filming and photography could be a satisfactory means of representation. In that case, reconstructing and completing gaps in heritage through the productivity and creativity of digital technologies place them as productions of the present, and not as re-creations of the past.

The numerous technological innovations which have taken place in the last two decades on computer graphics have had an immediate influence on experiencing and interpreting cultural heritage and on how it will be presented to the wider public. The general audience seems to easily consume pleasant wonderful pictures of virtual heritage and museums and heritage organisations tend to support the “sale” of the past covered in colourful and impressive shapes (Barceló 2002).

Virtual productions have been into a position where the original heritage itself becomes separated and different from the virtual mechanical products. Nevertheless, every kind of interpretation, virtual or not, bears a new facet different from the real event. By definition VR refers to a more fictional and imaginary environment and the concept of virtual heritage tends to present not the genuine in existence, but conversely the desire for the one that is missing.

The paradox of VR is that it is the main practice to depict heritage reconstructions when the original models are not available to duplicate; whether that is because the majority of them have been damaged or changed. Therefore, VR depicts reconstructions which do not exist as originals, giving rise to the creation of new images of the past so close to reality that they could be perceived as the ‘true’ realities. Baudrillard (1994) in his theory on Simulacra and Simulation sees cultural products as results of copies without archetypes. From his perspective contemporary societies have put, signs and symbols in the place of reality and he advances the notion that human experience is dominated by a simulation of reality which
has no references or foundations. It can be reasonably asserted that virtual heritage is a simulation without original as its main purpose is to depict what can no longer be seen and whose existence is ideal virtually through the use of signs and symbols.

Based on Baudrillard’s theory and in the general complex of virtual heritage reconstructions, VR depictions are simulations as copies without originals. His statement that the real will not have the chance ever again to produce itself confirmsthat computing designers do not produce accurate reconstructions of heritage sites but simulations of what they could have represented in the past. Hence, this confirms ‘the generation by models of a real without origin or reality’ and the tendency of virtual heritage to copy and reconstruct cultural sites without having an original.

Although VR could not represent the reality of the past as it has happened, it is still seen by the general public as a medium which can enable spectators to travel to the past and experience ancient sites and cultures. Its continuous interactive displays in museums and cultural organizations could be almost enough to confirm its contents. Consequently virtual heritage is presented as a correct reproduction of the “factual” and therefore it is absorbed with the tangible heritage in significance. The transformation of tangible heritage to electronic interpretations means that digital systems have been embodied in the value and consumption of heritage as an objective and accurate reproductions.

Nonetheless, what seems to draw users to virtual environments, as Roussou comments (2008, 230), is not so much about achieving realistic representation as it is about abstracting away from reality and creating a believable and convincing environment, regardless of whether the imagery emulates the physical properties of the real world or not. VR in heritage presents that which is more than the factual and the possible; it designs the imaginary and the ability of the objects to be more than what they are or used to be in their physical existence.

3. From Interpretative Tool to Artwork

While in the previous section, virtual reality in heritage was examined as an area generally unmapped from a critical viewpoint, this section explores virtual heritage, not only as a tool and a method for interpretation as it is usually represented, but in addition as a material expression in its own right with its own aesthetic qualities.

One of the basic characteristics that VR presents in multimedia technologies to an even greater degree than the traditional media such as film and photography, is the transformation from its material tangible state to intangible ways of experiencing culture. The shift has been to explore heritage from a close-up point of view, typically face to face in its physical presence to, today, where it is a disembodied and animated design. Objects have ceased to be the principal point for the experience and interpretation of heritage. In other words, it can be maintained that the value and core of heritage change their reference point from the object to the medium by emphasising the method of the experience over and above what is being displayed.

Witcomb (2007, 36) comments on the progressively increasing numbers of museums which use immersive technologies in their exhibitions, not as extra interactivities and multimedia, but as objects. According to her, the materiality of virtual technologies have come to stand as displays in their own right, almost replacing the historical role of objects in exhibition making. In a sense they have become objects. Thus, part of this concept is that the use of VR in exhibitions is perceived not merely as a tool to interpret the display contexts where visitors can get information through touch screens and Web sites but beyond to the use of an extra interactive stand.
VR has affected the presentation of collections in museum exhibitions and has come to stand as an artwork that can be equated with artefacts. That is for three reasons; firstly, virtual heritage products stand in galleries and can be viewed in three dimensional form, as almost every object can be; secondly they display artefacts from different angles and views and thirdly they are products of human creativity. In that sense, the practices of VR in heritage share basic characteristics with material heritage and can therefore affect the traditional displays of material based presentations in museums by shifting from being an interpretative tool to becoming the central object of an exhibition.

This direction of movement is brought about by the practices and methods of VR itself because what seems to be crucial for the user is not to face the artefact or site itself, but to access the desired information via a great range of experiences, engagements; even in the form of an entertainment; all of which can be offered by the new media. Travelling to heritage sites through new media technologies has become ‘a method of approach based on human realities rather than “culture” as Ronald Barthes states in *Mythologies* (1972, 76) where he analyses the consumption of mass cultural productions. As the value and consumption of digital systems have been embodied in the postmodern society, the need for VR and immersive technologies to understand and interpret history is as natural a consequence just as computer technologies are deemed necessary tools in the modern human life. Heritage is seen to be important as it is more accessible to the public when it contains and gives information that is communicated by a variety of new technologies such as VR.

Furthermore, it can be contended that what is experienced first of all in a virtual environment, is not the heritage but the multimedia technology. In the context of virtual 3D simulations, VR has the potential to become the contemporary symbol of the past and as a state-of-art medium bears the fear of being accepted as the new version of the obsolete object. Nevertheless, virtual heritage cannot be viewed as offering a completely different message from real heritage because it always carries its value within and refers directly to its presence in the history. As Barthes (1972, 5) claims ‘the photograph always carries its referent with itself...They are glued together limb for limb’. In virtual reality, without the prior existence of the real heritage, virtual simulations would not have a meaning or a reference.

Whilst such an approach sustains that virtual heritage can stand as independent reality when simultaneously bearing its reference, one could argue that virtual reconstructions do not threaten the ‘aura’ of the artwork as Benjamin describes. Since a whole creative process is being pursued in order to interpret virtually how heritage sites and artefacts were represented in the past, the new product is standing as an independent reality and not as a surrogate of another entity.

Thus, it could be suggested that the process of simulation, not only as the subject of virtual reality reconstructions but also as it is related to the general complex of heritage, could not be consumed as a monotonous, dry reproduction but as a productive piece of work, one that stems from human creativity.

To summarise, in seeking to identify the nature of VR and its presence in an exhibition context, the competitive relationship between the object and its virtual reconstruction should be noticed. By understanding virtual heritage as an independent creation that features the elements that an object consists of, means that it can be argued that from a state-of the art application it can become an artwork in its own right.

### 4. Creating New Heritage Virtually

The ‘Kyoto Virtual Time-Space’ (hereafter Virtual Kyoto), a research project developed by
The Department of Geography at Ritsumeikan University in Japan is an outstanding example of how virtual heritage sees the past from the perspective of the present and how technology can help to preserve and restore the heritage pipeline and be an important tool for informal education.

The main purpose of the project is to create a virtual city model from the late 8th to the 21st century. The advanced Geographic Information System\(^5\) which is used to create, on a large scale, the urban 3D model of modern Kyoto, reconstructs with great accuracy the modern and ancient city plan (Fig. 1) (Yano et al. 2007; Yano et al. 2009; Isoda et al. 2009, 21-22). The purpose of Virtual Kyoto is to create an online city model where visitors can travel virtually in the historical city (Figs 2 - 3) (Yano 2011). The 3D city planning starts with the reconstruction of modern Kyoto and goes backwards to the past by replacing the modern buildings with the older following the time period settings, Heisei era (1990s - present), Showa era (1920s – 1980s), Taisho era (1910s), Meiji era (late 19th – early 20th century), Edo era (17th – late 19th century) and finally Heian era (late 8th – 12th century).

In order to develop the Virtual Kyoto the following activities were undertaken. Firstly, material works were collected which

\(^5\)Geographic Information System (GIS) or Geospatial Information System is a programme created to merge geographical and cartographical data technologies.

---

**Figure 1.** Modern city plan of Kyoto in virtual reality.

**Figure 2.** Virtual streets of Kyoto.

**Figure 4.** 3D models of buildings in Kyoto.
Virtual Reality Simulations in Cultural Heritage
Ioanneta Vergi

are important references and documentations in terms of the topography of Kyoto through time, for instance contemporary and old topographic maps, aerial and streets photos, picture maps, significant landscape paintings, archaeological sites and historical information. Secondly, the 3D digital models of all existing modern buildings, traditional townhouses and historical monuments such as temples and shrines were created. Thirdly, the above data was transformed into virtual models based on the simulations and estimations of the project developers (Fig. 4).

Virtual Kyoto is created for professional use by city planners and architects and is accessible to the wider public via the official website of Virtual Kyoto. One of the characteristics of Virtual Kyoto is that it represents a four-dimensional space, in which a temporal dimension has been added to the three physical dimensions. This distinct element of the project places it in the category of four-dimensional city modelling, i.e. 3D plus a time dimension (Kyoto 2011). Virtual Kyoto has an important role in assisting urban landscape planning in Kyoto as well as wending rich information about the historical city to the world through the Internet. The innovative pedagogical value of the project is the development of the first digital archive of important tangible and intangible cultural heritage of the city of Kyoto, including scenery, paintings and festivals. The results are now being used in global-level research and as educational resources for young researchers. In the near future, the research team of the project intends to turn the programme into a digital museum which will be a meaningful contribution to the local community and wider society.

The project runs on a powerful computer database and visitors can explore Kyoto streets using their computer mouse. By choosing a timeline they can navigate through the city and examine how Kyoto looked at different times in history based upon historical information. When the visitor clicks a different time period, the presentation of the city changes to what it looked like during the period selected. The programme includes replications of every single street of the city with local Kyoto townhouses, known as Kyo-machiyas, modern houses, warehouses, shrines and temples. The project attempts to replicate from sounds, crowds to traditional festivals such as the Virtual Gion Festival one of the most popular festivals in Japan which take place every year in Kyoto.

Examining the product of Virtual Kyoto, it could be argued that it presents an innovative way to form new knowledge model for the city of Kyoto in line with contemporary pedagogical and theoretical public access concerns and the capabilities of emerging technological innovations in digital spaces. Virtual Kyoto illuminates the form and nature of new knowledge models in heritage practices, how these relate to users’ needs and abilities, and how changing paradigms can reform museum and heritage practices, such as in institutional knowledge creation and collection documentation.

5. Conclusions

New technologies can offer great capacities for the improvement and advancement of cultural experience. The acceptance of virtual reality as a method of interpretation and
display in cultural heritage could be presented on the one hand as threat; that is it might be understood as reality and on the other hand it could be seen as an innovative art of digital technology with its own creativity and quality features. The value of virtual reality in the heritage sector depends on its promotion and management by cultural producers.

Looking at the development and progress of earlier visual technologies, for instance photography, it could be envisaged that the future of virtual heritage will be new viewpoints for the presentation of artifacts in exhibition contexts and a revolution on how cultural producers can create a new heritage out of an older one. The installation of virtual reality technology in order to interpret cultural heritage and to mediate it to a wide public should not be seen only as a simulation in cyberspace but as a part of computing science history and ongoing development where its value is formed both by the ever-shifting cultural approaches, social mores and politics.

Acknowledgments

The author would like to thank Professor Keiji Yano and Dr Simon Kaner. This paper includes imagery generated using the Virtual Kyoto Time-Space project, which is a project of the Department of Geography at Ritsumeikan University in Japan.

References


Taking Excavation to a Virtual World: Importing Archaeological Spatial Data to Second Life and OpenSim

Isto Huvila  
Uppsala University, Sweden

Kari Uotila  
Muuritutkimus, Finland

Abstract:  
The benefits of analysing and presenting archaeological spatial data in an interactive 3D environment have been discussed extensively in the literature. This paper reports of a R&D project that explored the possibilities of presenting archaeological information in virtual worlds with a specific focus on presenting and using actual documentation data captured by total stations and laser scanners directly in virtual worlds. Field trials were conducted in Second Life and OpenSim environments. The findings indicated that Second Life was a preferred environment because of the relatively large existing ecology of individual and institutional users. The proprietary nature of the environment and the consequent limitations to data transfer and the control of the world did, however, make the importing, linkage and manipulation of data problematic. OpenSim allowed a total control of the environment, but lacked certain technical features implemented in Second Life together with a comparable, large population of users.

Keywords:  
Documentation, Second Life, Opensim, Spatial Data

1. Introduction

The benefits of analysing and presenting archaeological spatial data in an interactive 3D environment have been discussed extensively in the literature. Researchers have suggested using various techniques from proprietary visualisation technologies (e.g., Rua and Alvito 2011) to open 3D engines developed for first-person computer games (e.g., Anderson 2004) and approaches based on open standards such as VRML (Drap and Long 2001). Several authors have also recognised the benefits of social exploration of data in open online environments such as ActiveWorlds (e.g. Forte and Beltrami 2000; Prasolova-Forland et al. 2007).

The development of commercial and open non-game 3D virtual worlds has been rapid since the turn of the millennium. Even if the most intensive hype around worlds like Second Life has been relented and some of the pioneers including There.com have been closed, multi-user virtual worlds have shown their usability in several areas including education (Holmberg and Huvila 2008), cooperative work (Smith et al. 2009), certain leisurely contexts (Crowe and Bradford 2006) and, for instance, archaeology (Graham 2007).

This paper reports of a research project that explored the premises of presenting archaeological information in virtual worlds with a specific focus on the affordances and constraints of presenting and using actual spatial data captured by total stations and laser scanners directly in virtual worlds. Field trials were conducted in Second Life and OpenSim environments. The project tested different types and forms of spatial data from several different sources and evaluated approaches for the presentation of the data in the virtual world environment. There is some earlier research on the archaeological applications of Second Life even if most of the published studies tend to focus on individual archaeological exhibits.
In contrast, OpenSim has been discussed to a lesser extent (besides the suggestions for future work by Forte and Kurillo 2010), and upon our knowledge no earlier comparative studies exist.

2. Second Life and OpenSim in Archaeology

Second Life has been used for the presentation of archaeological projects and sites since the mid 2000s (Graham 2007). Potential archaeological applications of virtual worlds have been discussed in the context of a number of different platforms including ActiveWorlds (Prasolova-Forland et al. 2007), Blue Mars (Ober 2012) and There.com (Pearce 2009), but there is no doubt that Second Life has been the most popular environment. Forte and Kurillo (2010) make some remarks on the possible future directions of their research and the possibility of using OpenSim for popular presentation of archaeological sites and Open Cobalt for visualising archaeological datasets, but otherwise very little has been published on the archaeological uses of that particular platform. Much of the work has focused on the promotion of archaeological and heritage sites (e.g. Chávez-Aguayo 2011). Virtual worlds have been conceptualised often as a primarily representational technology (e.g. Grindley 2007). Museums and archaeological projects have built replicas of archaeological monuments and “visitor centres” for dissemination of information on on-going projects and sites of interests (e.g. Getchell et al. 2009).

Besides the promotional and entertainment use, one of the major areas of the use of Second Life has been education. A large number of educational institutions have participated in Second Life, and even if the general interest has been in slight decline since the peak of the hype in 2006/2007, an extensive corpus of research contains plenty of evidence of the benefits of using virtual worlds for educational purposes (e.g., Holmberg and Huvila 2008; Olasoji and Henderson-Begg 2011; Sköld 2012). In context of archaeology, for instance, Getchell et al. (2009), Graham (2007) and Salmon et al. (2010) describe how Second Life has been used for educational purposes. Students have been given a possibility to visit virtual archaeological sites (Salmon et al. 2010) and re-enactments of cultural environments and to, for instance, curate exhibitions (Getchell et al. 2009). From a scholarly perspective, one of the most ambitious attempts to integrate Second Life into the frameworks of public archaeology and scholarly archaeological processes have been conducted in the context of the Çatalhöyük project (Tringham 2010). Morgan (2009) has suggested that from the perspective of archaeological scholarship, the most significant benefit of virtual worlds might be their capability to provide archaeologists an easy-to-use sandbox for challenging the “static modes of representation” and at the same time to provide opportunities for non-expert participants and audiences to access archaeological sites and knowledge.

Even if the most of the authors have been rather optimistic about the possibilities of using Second Life and other virtual world environments in archaeology, Harrison (2009) reminds that it is important to consider the consequences of the virtualisation of cultural heritage. Even if the notion of cyber-archaeology of Harrison may be seen as conceptually somewhat problematic term, the necessity of meta-research on the use and consequences of virtual worlds is similarly apparent in the field of archaeology than it is in the context of cultural heritage studies.

3. Archaeology in Virtual Worlds (ArVi)

The present study was conducted under the auspices of the Archaeology in Virtual Worlds (ArVi) project financed by the Finnish Board of Education in 2009-2011. The aim of the project was to test the portability of archaeological research and documentation data and its usability in open three-dimensional
virtual environments. The project used Second Life as a primary case example, but evaluated also other platforms. In addition to technical testing, the project looked into how Second Life and other comparable virtual worlds could function as platforms for communicating archaeological information online and as a part of a physically based exhibition. The reason for focusing on open (in the sense of not being restricted for particular institutions and users) virtual worlds was to evaluate the possibilities to take archaeology to virtual environments with an existing user base. The two chosen platforms are relatively stable and provide existing global 'standard' environments for visualisation and exploration of three-dimensional information.

Second Life was chosen as a test platform because of its popularity, the relative openness of the environment for developing and presenting user generated content and because of the earlier positive experiences of its usability as a learning environment (Holmberg and Huvila 2008).

The project was conducted as a cooperative effort between the municipality of Eura and Muuritutkimus Co., a private archaeology consultancy. The research was conducted by Dr Isto Huvila (Uppsala University) and Dr Kari Uotila (Muuritutkimus) with contributions from Mr Markus Kivistö, Ms Tuija Väisänen and Mr Pekka Mäkitasku.

4. Findings and Observations

A significant reason for the popularity of Second Life is undoubtedly the availability of simple and relatively powerful 3D modelling tools directly in the freely available client software. In spite of their virtues, from an archaeological point of view, the tools have certain limitations. The built-in instruments are based on primitives based modelling paradigm and are very simplistic in comparison to professional 3D packages. The processing of 3D data is further complicated by the fact that the number of available polygons in each area of the virtual world is limited because of performance restrictions of the environment. Due to the limitations, it is often preferable to use simple geometry and, as much as possible, to simulate details with texture maps.

At the time when the project started the possibilities to import 3D data into Second Life were severely limited. It was possible to import specifically prepared bitmaps (sculpt maps) for 'sculpting' meshes inside the virtual environment. In addition it was possible to generate and modify objects by using scripts produced by add-ons to external modelling software such as Blender or 3DS Max. In late 2010 a new feature was added to the Second Life software that allowed users to upload meshes directly into the environment in Collada format (www.collada.org). The technique was implemented almost simultaneously to OpenSim environment.

In the ArVi project, the importing of models was tested in both Second Life and OpenSim environments using both primitive and mesh based approaches. The tests were conducted using digitised (scanned and photographed) and born-digital documentation data from the iron age archaeological site of Luistari (located in Eura, south-western Finland), and later on using additional laser scanning and CAD-based documentation data from an excavation conducted in 2011 in the World Heritage site of Old Rauma. Laser-scanning data from Faro Focus 3D was prepared using Faro SCENE 4.8 software and combined with total station measurement data and processed further in AutoCAD Civil 3D version 2012. Both types of data were imported into MeshLab version 1.3.0a for post-processing and conversion to Collada format.

The primitives based approaches were found (as expected) to be largely useless for processing existing documentation data. It was apparent from the beginning that the tools are mainly usable for creating new models. The
Taking Excavation to a Virtual World: Importing Archaeological Spatial Data to Second Life and OpenSim
Isto Huviila and Kari Uotila

The technical process of preparing and exporting documentation data into Collada format and to Second Life and OpenSim was mostly a relatively straightforward task without major problems. The main limitation related to the size of importable meshes and to the need to keep the level of details very low. In practice, it was apparent that for presentation purposes, the direct use of actual documentation data was possible only for visualisation of individual small objects and details. Because of the need to keep the level of details low, it was necessary was necessary to create the main part of any site and landscape using the primitives based modelling tools and bitmap based landscaping of the environment. The tests confirmed the apparent fact that both the embedded modelling tools and the mesh import functionality have not been developed for working with typical archaeological documentation data, but to help 3D designers to develop functional and aesthetic 3D objects and environments from the scratch. In practice, our tests showed that it is relatively unproblematic to import files containing up to 60,000 vertices even if the particular figure is not a definite technical limit. The mesh and scene scales were also limited and occasional stability issues occurred. Even if the OpenSim documentation states that the mesh support is under development and not necessarily 100% Second Life compatible, in our tests OpenSim tended to be more stable than Second Life. The practical tests conducted as a part of the ArVi project confirmed also the possibility to upload and use typically slightly larger and more complex models to OpenSim test environment than to Second Life.

On the basis of the evaluation of the usability of the environments, it could be stated that Second Life had an advantage of providing a relatively large existing ecology of individual and institutional users. This advantage is especially apparent when compared to closed, stand-alone environments (e.g., as in Rua and Alvito, 2011). Second Life provided the best possibilities to take archaeology to an environment with a relatively large ‘native population’. At the same time, however, the proprietary nature of the environment and the consequent limitations to data transfer and the control of the world made the importing, linkage and manipulation of data problematic. OpenSim allowed a total control of the environment, but lacked certain technical features implemented in Second Life and a similarly large existing population of users than Second Life.

Second Life (Fig. 1) provides a broad array of tools for presenting and visualising archaeological excavations. In addition to displaying a three-dimensional model, it is possible to use text, audio files and video within the three-dimensional environment. It is also possible to create links between the Web and the virtual environment and to reuse and build on the data that is already available online both in popular ‘web exhibitions’ and open documentation systems and repositories. The virtual environments provide the general public also with a possibility for personal engagement with experts. Archaeologists can organise guided tours in a virtual excavation site and interact directly with the public without a need to travel to a specific physical location. The virtual excavation site can also be used as a site for public programming such as for giving lectures and organising other types of events. A similarly significant aspect of the environment is that the public can engage with the virtual excavation at the time of their choice, an aspect that is seldom possible at an authentic excavation site. From the perspective of the physically based
presentation of archaeological sites, the virtual presentation is capable of creating added value as a resource that gets people interested in the site, provides additional in-depth information, possibilities for augmented experiences and modes of interaction that could compromise the preservation of the authentic site.

Technically, in order to be able to interact in the virtual environment, each particant (or resident, using the common Second Life designation of its users) needs to register a free personal Second Life account, to create an avatar and to download and install Second Life client application. The terms of service of Second Life limit the participation for individuals of over 16 years of age. In practice, an adult can show Second Life environment for younger individuals using his or her own avatar, but the experience of watching others navigate in the environment lacks authenticity and engagement. It is obvious that a typical Second Life user is not necessarily a frequent visitor to archaeological sites, but it is reasonable to assume that the original lack of interest may be compensated by their capabilities to interact in the Second Life environment. A Second Life user is likely to have necessary technical skills for navigating in the environment and engaging in a meaningful interaction with a virtual archaeological site. The learning curve and the need to have a Second Life account limits the usability of the environment in traditional exhibition environment, but does not entirely rule out a possibility to use it as a complement.

One of the challenges of operating in the commercial Second Life environment was experienced in the end of 2010 when the Linden Research, the creators of Second Life, announced the discontinuation of their educational and non-profit pricing scheme of virtual land. Another economic shortcoming of the Second Life environment is the cost of uploading image data into the virtual world. Even if the cost of uploading a single unit of data is low, it became apparent that especially in the test phase when a lot of materials need to be uploaded for testing, costs could be significant even if they are unlikely to become prohibitive if the most of the testing will be conducted in a separate gratis test grid (i.e. world). Another, potentially a more significant issue for the project was the limited possibilities to export data from the virtual world to other environments.

OpenSim provides a largely Second Life compatible open source server environment that can be accessed (similarly to Second Life), using the Linden Lab Second Life viewer software and an array of independent virtual world viewers. The two platforms are visually similar to each other. Even if the two environments are similar and compatible on a protocol level, Second Life is more mature and stable environment with more features. By the time of this writing, it is, however, necessarily to remark that OpenSim is being developed in a fast pace. The possibilities to install a local instance of the OpenSim software or to choose a hosting service provider from a relatively long list of companies and other organisations provide flexibility and better possibilities to control the platform. The use of OpenSim environment eliminates also the need to comply with Second Life age limits and a need to use a personal avatar. In this sense, OpenSim is a more flexible solution for public access terminals. At the same time, however, operating an OpenSim based world does potentially require slightly more technological expertise either in terms of installing and maintaining own server or in terms of selecting.
the best possible host organisation. Another more significant shortcoming in comparison to Second Life is the loss of a relatively large virtual world with a population of tens of thousands of users. Even if the OpenSim worlds can be connected to each other into similar grids (i.e. worlds) than Second Life and it is possible to travel between individual grids using so called Hypergrid technology, the user base of OpenSim based worlds is significantly lower than that of Second Life. In the present study, this issue was judged to be of particular significance because the focus of the project was to evaluate open virtual worlds as an alternative to closed ones.

5. Conclusions

In spite of the several technical challenges identified during the process, the virtual worlds provide a highly promising environment for presenting archaeological information using authentic archaeological data. The earlier archaeological projects in virtual worlds such as Second Life have been based on using the environment primarily as an easy-to-use exploratory modelling tool for archaeologists (e.g. Morgan 2009) and a showcase for purposely-built visual re-enactments of sites and monuments (e.g. Getchell et al. 2009). In spite of the presence of certain challenges and limitations, the present study was able to pinpoint several possibilities and benefits of working with real documentation data in these environments.

It is apparent that none of the tested environments was entirely ideal for the evaluated purposes. Second Life had the largest population of existing users, but posed limitations to the level of (technical) control of the environment. The relatively small user base, the need of slightly higher technical skills and the lack of the maturity of the environment limited the usability of OpenSim. It is necessary to note, however, that the fast pace of the development of OpenSim and the emergence of new OpenSim hosts is likely to decrease the impact of the last two issues. At the moment, the selection of a virtual world is a trade-off between available features. If the existing user-base would be the only criteria of significance, an ideal choice at the moment would be an environment like the World of Warcraft or Habbo Hotel, but it is apparent that they lack features that would make them ideal for presenting archaeology, especially, using authentic three-dimensional documentation data.

References


Using ConML to Visualize the Main Historical Monuments of Crete

Panagiotis Parthenios
Technical University of Crete, Greece

Abstract:
Conceptual models offer the ability to capture several concepts, and more importantly their often complicated relationships, in one single view. Especially when dealing with complex and deep hierarchies or intangible notions, a conceptual model can offer an additional level of perceptual understanding. We use the Conceptual Modelling Language (ConML) in our proposed application for the presentation of the main monuments of Crete as a tool for organizing, manipulating, and communicating the large amounts of data such a project entails. Conceptualization and abstraction of information through different levels of detail allows the application to be light and easy to use. Moreover, the ability to switch between different historical periods offers a comparative study of the monuments evolution in time.

Keywords:
Conceptual Modelling, 3D, Abstraction, Visualization, Information

1. Introduction

Crete is the largest island of Greece, located to the south, famous for its rich cultural history, which dates back to the Middle Paleolithic age, 128,000 BC. Crete was the center of the Minoan civilization (2,700-1,420 BC). Since then, a large number of monuments has been documented throughout the different historical periods, the most important of which are the following seven (7):

- a. Minoan
- b. Hellenistic
- c. Roman
- d. Byzantine
- e. Venetian
- f. Ottoman
- g. Modern

Our goal is to design an online platform open to the public for the promotion of the cultural heritage of Crete, through a simple, user-friendly intuitive environment. Our prime challenge has been how to manage such a large amount of information over the internet, in a transparent, light and simple way for the end user, in addition to offering the ability to compare data over time, during the historical periods. In order to achieve this we have been using the notion of Conceptual Modelling along with the principles of Model Based Information and Object Oriented Databases. The idea is simple: instead of having all information to its full extend available up front, we break it into nodes, levels of abstraction, called “Levels of Detail”, providing the minimum information needed at each given time. Information is stored on each object, each monument, along with its different Levels of Detail. The Levels of Detail that we are using in this platform are the following five (5):

- a. Prefectures
- b. Cultural provinces
- c. Settlements - Towns
- d. Building complexes
- e. Buildings - Monuments

The conceptual diagram on which this platform is based can be seen on Figure 1.

2. The Platform

The platform is comprised of the main, central space, where the 3D models are presented, and two scrollable sidebars, one horizontal and one vertical. The horizontal one controls time, and allows the user to switch between the seven historical periods and the vertical one controls the level of detail,

Corresponding author: parthenios@gmail.com
allowing the user to switch between more or fewer abstract modes. At the same time, the user has the ability to navigate in real time in the main space around the models, using pan, zoom in/out, rotate, etc. Only when the user reaches the fifth Level of Detail, that of a single monument, he/she has access to all the available related information, which depending on the type of monument could be:

- Photographs,
- Architectural drawings (of the existing and/or the restored monument),
- 3D model (of the existing and/or the restored monument),
- Walkthrough animation (of the existing and/or the restored monument),
- Video of the area as it is today,
- Maps of the area,
- Related documents with more information,
- Related links to other websites,
- Keywords. Use of keywords allows a cross-reference function independent of the 3D models.

The philosophy of the monument presentation is intended to address mainly non-experts, therefore it follows a more abstract and simplified view of information. It should be easy to use for a visitor who does not have a deep knowledge about Crete and its civilization and would like to be informed at a glance what to visit and where. On a second level, the visitor can focus more on a group of monuments and prepare for his/her visit acquiring more specific information, stored “on” the monument’s model itself. The application is currently based on Adobe Flash in order to provide maximum compatibility with most of the major web browsers, to be light and easy to use and to avoid installation of other software. At the same time we are investigating whether technologies such as Unity 3D or SpiderGL can provide as a more suitable environment to work with.

The proposed application could take advantage of other related research projects which have rigorously documented and categorized the monuments of Crete, such as the “Digital Crete: Mediterranean Cultural Itineraries” (http://digitalcrete.ims.forth.gr), which was implemented under the framework of the Greek Operational Program Information Society (Action 1: Education and Culture, Measure 1.3: Documentation, Management & Promotion of Greek Cultural Heritage) (http://www.infosociety.gr).

3. Using the Conceptual Modelling Language (ConML³)

There are a number of languages suitable for conceptual modelling, such as CIDOC CRM or UML. The reason we chose ConML (Gonzalez-Perez, Parcero-Oubiña 2011) is because ConML is easy to be utilized by non-experts in information technologies, is simple and can prove to be expressive in complex domains such as those in the humanities. In order to begin
building our conceptual model, first we have to define our main classes: the class “Object of Interest” and the class “Representation”. The fundamental argument on which our model is based is: “every Object of Interest is represented through a Representation”. The Objects of Interest can be one of the following five (5) classes, which are Subclasses of the “Object of Interest” class: “Monuments”, which can be part of a “Building Complex”, which can be part of a “Settlement”, which can be part of a “Cultural Province”, which can be part of a “Prefecture”. Each of these five (5) Objects of Interest can have attributes, such as “Description”, “Links” and “Keywords”. They must all have a common attribute though, the “Historical Period”, which therefore becomes an attribute of their abstract class, the “Object of Interest”. The data type of the attribute “Historical Period” is enumerated type and can take a value of one of the following seven (7): Minoan, Hellenistic, Roman, Byzantine, Venetian, Ottoman and Modern. Since one Object of Interest has to belong to at least one Historical Period, but could also belong to more than one, the cardinality of “1…*” is placed next to the name of the attribute. The class “Representation” has the following subclasses: “Photo”, “Drawing”, “Video”, “Map”, “3d model”, which have various attributes such as “Exterior”, “Interior”, “Resolution x”, “Resolution y”, “Color”, “View”, “Reality”, according to their type. Their common attributes become attributes of their abstract class, the “Representation”, and are the following: “Analog” (type: boolean), “Copyrights” (type: text), “Year” (type: number), “Description” (type: text) and “File Format” (type: enumerated).

### 4. Conclusions

The primary contribution of the proposed platform is the ability to capture time in a comparative format. Nevertheless, the fourth dimension is exploited here in a more abstract way than other scientific approaches (Kulitz, Ferschin & Matejowsky 2010), (Mylopoulos 1992) since realism and full detailing is not the goal of this application. Furthermore, its advantages are: the user friendly interface which is addressed towards non experts and its ability to continuously expand with new material regarding either new monuments or new information for existing monuments. Some of the issues we are currently working on are:

- Subjectivity due to abstraction. When information is abstracted, the role of the person who decides which information should be secluded is a key role since it could possibly skew the end result.
- Uncertainty due to lack of information.
- Use of multiple semantic links.
- Interoperability / expansion. The proposed application could serve as a central platform which could be joined by other applications which focus on a more detailed, photorealistic monument representation.
- Building Information Modelling (BIM). What can we learn from the structure of information used today in BIM?

### References


A High-Performance Computing Simulation of an Irrigation Management System: The Hohokam Water Management Simulation II

John T. Murphy
Argonne National Laboratory, USA

Abstract:
The Hohokam Water Management Simulation II is a reimplementation of a desktop-based simulation of the operation of a large-scale irrigation system for a high-performance computing (HPC) environment. The focus of the simulation is the examination of the Hohokam, who lived along the Salt and Gila Rivers in the US Southwest and who constructed and maintained a large-scale irrigation system that was in use for nearly a millennium. The new simulation implementation is specifically designed for the Blue Gene/P supercomputer (BG/P) maintained by Argonne National Laboratory (ANL), which can perform operations on more than 100,000 processors simultaneously. Work on the simulation framework is ongoing, with the goals of using the highly parallel processing environment of the BG/P in two ways: first, by executing single simulation runs across multiple processes, making possible both simulations at larger spatial scales and those at finer chronological and spatial resolution; second, by permitting large numbers of concurrent runs to be executed strategically to explore the extremely large parameter space made possible by various combinations of input data, parameters, spatial and temporal scales, and algorithms available to the simulation. Details and challenges of this process are described.

Keywords:
High-performance Computing, Agent-based Modelling, Hohokam, Hohokam Water Management

1. The Hohokam Context

   The Hohokam lived in what is today the southern part of the U.S. state of Arizona, a challenging desert environment but one in which they flourished for over a millennium. They are best known for constructing large-scale irrigation works, the most dramatic of which stretched from the Salt and Gila Rivers in the area of what is today Phoenix, while others at smaller scales drew water away from various sources to the south. The temporal and spatial boundaries of ‘Hohokam’ are often open to debate (see Bayman [2001] for an overview), but the florescence of the large-scale irrigation works took place beginning around 400 AD. These were operated successfully for nearly 1,000 years, before being essentially abandoned shortly before the arrival of Europeans to the New World. Partway through this trajectory, a dramatic change took place in the Hohokam way of life; around 1150 AD, settlement patterns, domestic and public architecture, and rituals that had been in place for six centuries or more were replaced with new forms that suggested that the set of personal, kinship, and community relationships among the Hohokam were being changed. The means of operation of the canal systems before and after these changes are the objects of ongoing investigation (for example, Woodson 2010), including a simulation modelling effort (Murphy 2009, 2012) that will be the focus here.

1.1 Challenges and Opportunities

   In the attempt both to model and to understand the Hohokam system we are faced with a number of challenges. These begin with the data: although a large number of datasets exist that shape our picture of the Hohokam, these are often incomplete, or at least not complete enough to support a modelling effort, and they are of various kinds and thus are not easy to integrate into a coherent whole. They inform an understanding of the Hohokam that, reflecting the disparate data, exists at multiple scales and is open to multiple interpretations.
indeed, researchers take widely varying approaches in offering explanations of the Hohokam trajectory. The boundaries of what is to be explained are, likewise, open, and the idea that all the elements within these boundaries form part of a coherent system cannot be taken as given and instead must be a subject for investigation.

However, contexts such as the Hohokam case offer an array of opportunities that make it enticing to overcome these challenges. Irrigation management has long been of interest to anthropologist, archaeologists, and those interested in the dynamics by which larger political units (such as states) may form. The Hohokam may represent a system in which principles of self-organization led to a condition in which the overall system was run efficiently without the need for a central authority, similar to the example offered in Bali by Lansing and Kremer (1993). The change in the character of Hohokam society and its eventual decline present additional questions that can be examined in terms of new frameworks such as resilience (Hegmon et al. 2008). Perhaps the most ambitious hope is that the Hohokam can be shown to be a special case of a more general dynamic, one that is observable in other archaeological contexts and potentially useful in contemporary situations.

The initial modelling efforts that have been undertaken on the Hohokam case offer some examples of the possibilities to be found in exploring the Hohokam context; these have been presented in earlier work (Murphy 2009, 2012). The general lesson of these examples is that an approach that works from small assumptions and a few elements with a moderate degree of uncertainty can lead to complex and unexpected results.

This implies that the opportunities offered by modelling an archaeological system like the Hohokam case are wide. However, there are difficult problems to be solved in charting the course forward through the numerous possible modelling avenues available. Previously (Murphy 2007, 2009, 2012) it has been suggested that a new modelling strategy would be required for approaching problems such as those offered by the Hohokam context, and that this should be an ‘exploratory’ approach. By this is meant that the modelling effort should allow the research to shift the scale at which investigation takes place; it should allow the model to move along an axis of complexity and simplicity, removing complexity when it can be shown to be unnecessary; and to allow for the use of data in which there are varying levels of confidence, including elements that represent data attested from the archaeological record but also inferred, proposed, plausible, speculative, or purely hypothetical elements that may nevertheless lead to useful insight. This approach will require careful theoretical consideration, but also much more computational power.

2. High-Performance Computing

High-performance computing (HPC) is a field that is continually moving forward. The move toward greater computing power primarily takes the form of increasing the number of processors that participate in highly parallelized supercomputing systems. One such system exists at Argonne National Laboratory; known as the Blue Gene/P, it includes 40,960 computing cores, each one a ‘quad-core’ processor, allowing for a single job to be run across $4 \times 40,960 = 163,840$ processors. Each core has access to 2 GB of RAM (hence 80 TB for the entire machine). Communication among the nodes is accomplished by three separate networks for different patterns of interprocess communication. Additionally, Argonne has a parallel virtual file system that can accommodate very large quantities of data being generated by this system and allow it to be written efficiently while the job is running, and a visualization system that allows even extremely large sets of output data to be rendered for display and manipulated in real-time.
2.1 Repast HPC: An Agent-Based Modelling Toolkit for High-Performance Computers

The Argonne Blue Gene/P is impressive hardware, but it cannot be used for archaeological simulation without appropriate software. Software for archaeological simulation modelling at scale is only now being developed. At Argonne an in-house tool called Repast HPC is being developed that will serve as a general HPC simulation platform, and specifically will allow agent-based models to be run on high-performance computing platforms such as the Blue Gene/P. The hope is that this can form a starting point for a more general archaeological simulation toolkit.

Repast HPC (Collier and North 2011) is a general toolkit for what is termed ‘agent-based modelling.’ Agent-based models now have a comparatively long pedigree, but still represent a new and distinctive approach in social science modelling. The conceit of an agent-based model is that numerous instances of simple software agents are permitted to interact and follow simple rules; often these rules are shared by all the agents, though heterogeneous populations can easily be created. From their interactions global patterns can be observed that were not encoded in the agent-level rules, but arise- or ‘emerge’- from them, often in counterintuitive ways. The analogy to the way the real world works- comprised of independent actors but yielding population-level regularities- makes agent-based modelling a natural fit for the social sciences.

The parallel to Repast HPC in the single-processor (desktop) world is Repast Simphony (see http://repast.sourceforge.net), which offers modelers a number of convenient pre-packaged functions for defining agents, creating populations of agents, managing suites of simulations, and other common tasks. Repast HPC is written in C++ (vs. Simphony’s Java), and implements the core functions of Repast Simphony in a parallelizable way suitable for HPC environments.

A core component of the functionality of both versions of Repast is the idea of a ‘projection’. A projection is a means to structure agent relationships. Projections come in two flavors. Spatial projections arrange agents in some kind of space; 2-D grid spaces are commonly used, but continuous spaces and space in higher dimensions are possible. This imposes a logic on the agents’ interactions, so that some agents may be near to or far from other agents in a systematic way. The alternative is a ‘network’ projection, in which agents can be related to one another by arbitrary links in any pattern.

Repast HPC parallelizes an agent-based simulation by assuming that all agents have a ‘home’ process on which they are currently being managed. If information about an agent is needed on a process other than its home, a copy is made and shared with the non-local process. These copies, however, are passive: changes should not be made to the copy, and any changes made are not propagated back to the original. For spatial projections, the virtual ‘space’ is divided among the processors. Agents exist in particular regions of the virtual space and are at ‘home’ on the process that manages that region; agents that move across boundaries are moved from one process to another automatically. Interaction is presumed to occur only between agents that are close to one another spatially; for agents that are at the boundaries between processes, Repast HPC copies ‘buffer zones’ that will include copies of the agents within a limited distance on the far sides of the boundary. For network projections, agents on separate processes that are connected by a network link are mutually copied back and forth across the link, so that each process has the local agent and the non-local agent to which it is connected.

Although Repast HPC is quite powerful, it is not yet a mature ABM platform. Its initial releases (it is now in version 1.0.1- see http://repast.sourceforge.net) perform nicely for network-based or space-based simulations.
However, the Repast Simphony toolkit permits spaces and networks to be used simultaneously; in fact, any combination of spaces and networks should be permissible. Repast HPC’s ability to manage simulations with spatial and network components is limited- a serious constraint on Repast HPC’s flexibility.

Work is underway to address this; code now under development at Argonne will permit the same flexibility as Repast Simphony to mix spatial and network projections arbitrarily. Although much of the original Repast HPC architecture can be used in this effort, some larger changes have also been required; the anticipated date for release of this new code is unknown.

3. Archaeological Simulation in an HPC Context

The Hohokam simulation in its original version is an agent-based model; agents representing individual decision makers along the canal system (farmers, irrigation managers) interact with each other and with the virtual canal system (rivers, channels, fields, and crops). Hence it is natural to use Repast HPC to implement the parallelized version of the new simulation. However, the effort to do this has been slowed by the reconfiguration of Repast HPC that is currently underway.

However, enough work has been done to discuss some next steps, and address the larger theoretical question of how parallelized simulation can be applied to archaeological modelling. There are at least two distinct paths. The first is the division of a single simulation into separable parts that can be run in parallel. Repast HPC offers a good approach for straightforward cases. There are, however, variations that can create additional complications. In some cases, what is conceptually a single agent must be split and parallelized across processes. The Hohokam simulation offers an example: one variant of water flow calculations require values from all components of the canal system- hence across the entire simulation space; but this calculation can itself be parallelized.

The second area in which parallelization can be harnessed is in the exploration of larger parameter spaces. Here the computing power of the HPC platform is harnessed in a different way: instead of addressing a single simulation on thousands of processes, thousands of simulations are simultaneously performed. The versions of the simulations are selected based on their ability to inform some larger problem- they may, for example, represent the curve obtained when an input parameter is varied from the low to the high end of its possible range. With the ability to perform many such simulations in a short amount of time, the possibility to explore very large numbers of variant simulations is opened.

These opportunities, however, are not without challenges. As mentioned above, the development of the software needed to facilitate large-scale archaeological simulation on HPC systems is still underway. There are a number of problems that are or will be encountered before HPC systems can be easily employed in archaeological service.

One such problem is data management. This is a general problem in all of HPC, and is largely a technical issue, but it is one that any large scale archaeological simulation, particularly any that rely on large data sets of disparate kinds, and that generate the same, will have to consider.

A second, more theoretical question has to do with the specification of the axes along which the ‘exploration’ will take place. Work is currently underway to allow this to be done in an intuitive and useful way. The array of possible variations on a simulation can be immense; moreover, the possibilities need not form a ‘natural’ order. Whereas in earlier simulation approaches a manual strategy was
sufficient, at larger scales there will be a need have this exploration performed automatically. A schema that can allow combinations of variants to be defined in clusters and arranged in a meaningful (useful and interesting) order is key to exploring such a space.

A part of this will require a better definition of a ‘search strategy’: specifically, for a given axis along which a search is being performed, what constitutes ‘meaningful’ or ‘interesting’ results? How can it be determined that it is productive to continue to explore further or in greater detail in one region of the parameter space, and not in others?

If as was proposed above these simulations can take place at multiple scales, then it may be the case that the best scale at which to perform the simulation- that is, the number of processors that can perform the simulation, or sets of simulations, the most efficiently- is not initially known. Most HPC software now assumes that the scale of the simulation and the number of processes on which it will be performed is known before the simulation is run; software to shift scales on-the-fly is not available.

An additional concern is visualization. Simulations run in HPC environments lose the intuitive interfaces provided by desktop frameworks like Repast Simphony, yet visualization is key to understanding the dynamics of the simulation (especially in the early stages of work). Visualization tools for large-scale archaeological simulations do not exist. Further, the visualization of simulation dynamics is only one part of the problem: another can be visualization of the simulation search space and its results.

To all these concerns another can be added: the general unfamiliarity with social science that is found in the HPC world. As noted above, most HPC systems are put to use on very large mathematical problems; the idea of running thousands of small simulations is less familiar, and care sometimes must be taken to show that the power of the HPC system is warranted and represents the only way to approach the problem under study.

Work is underway on all of these fronts. Repast HPC examples exist that address data storage for agent-based simulations with millions of agents, and visualization of tens of thousands of agents (see Murphy 2011, Murphy et al. in prep). More directly, the framework that was used for the desktop version of the Hohokam simulation, called the ABCM framework (see Murphy 2009, 2012) is being implemented in a parallelized version; it brings along some strategies for exploring a complex parameter space, and, though more are needed, this supports the argument that the multiple simultaneous simulations in fact form a single problem that can only be approached using HPC. Part of the new framework is almost certain to make use of and be integrated with Repast HPC, and will be made available to the wider modelling community. Writing effective programs for HPC environments is difficult, but work to fully implement the ABCM framework and make it usable at very large scales is underway.

4. Computational Social Science

An optimistic view is warranted: the challenges just described will be overcome. Scanning ahead, then, some attention should be paid to the larger question of whether this increase in computational power will change archaeological simulation qualitatively, and, in the affirmative case, how. The larger field of computational social science is one that is rapidly shifting: data are now available on so many things, and in such great quantities, that new windows seem constantly to be opening into computational approaches to the study of human behavior. Archaeology, while perhaps not as directly in the flow of this torrent of data (the Hohokam have not left us ‘status updates’), will take a place in this. I believe that the key issue
that will arise when we can apply very large scale computing power to archaeological contexts will be the ability to consider our subjects computationally- that is, not only through the means of computational simulation, but as an example of a computational system. We can use HPC to think of the Hohokam as solving a problem- that of irrigation and cooperation in a challenging environment- and ask how they computed their answers.

Acknowledgements

This research used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357. This work was supported by American Recovery and Reinvestment Act (ARRA) funding through the Office of Advanced Scientific Computing Research, Office of Science, U.S. Dept. of Energy, under Contract DE-AC02-06CH11357. The original HWM Project was funded under Arizona State University’s Global Institute of Sustainability, with additional funding from the McDonnell Foundation. Special thanks are due to Steve Lansing, Ann Kinzig, and Charles Redman, and to the Repast HPC Development Team: Michael North, Nick Collier, Jonathan Ozik and the rest of the DIS CAS2 Team at Argonne National Laboratory. All errors are my own.

References


Field and Lab Recording
Application of RTI in Museum Conservation

Eleni Kotoula
University of Southampton, UK

Abstract:
This paper examines applications of Reflectance Transformation Imaging (RTI) in museum conservation and its contribution to conservation objectives. Recent developments in RTI technology have lead to a growing number of applications of RTI as a tool for the examination and analysis of surface detail and documentation, as well as a means of communication, dissemination and presentation. Applications of RTI to a wide variety of material and artefact types, as well as a wide range of conservation states, indicates that it reveals hidden details and provides an alternative solution to preventive conservation needs. Successful experimentation with micro domes, suitable for capturing microscopic RTI data, not only meets conservation needs but also signals a development that can broaden its application in the study of artefacts, particularly in museum conservation practice. The paper includes an overview of RTI and archaeological materials, as well as a pilot study for conservation documentation and cleaning monitoring.

Keywords:
RTI, Polynomial Texture Maps, Imaging, Conservation

1. Introduction

Conservation intervention is a never-ending process, which takes place after discovery and before the artefact’s interpretation. The changes introduced during this period, not only in appearance (including geometry, colour and texture), but also in chemical structure, are among the most influential processes that determine the artefact’s future, and dramatically affect interpretation and display. Nevertheless, conservation remains the least known period of the artefact’s contemporary treatment, not only for the general public, but also for cultural heritage professionals. Conservation dilemmas, technological needs and ethical issues form a delicate balance. Digital technology can play the role of a catalyst, and enhance inadequate traditional approaches with improved digital analogues, and by providing useful and affordable tools which bridge the gap between scientific developments and practical solutions.

Although research is based on a relatively long list of three-dimensional techniques for recording and visualization, as well as a number of computational photography and photogrammetric tools, for the purposes of the present paper attention is focused on RTI (Mudge et al. 2005) and one of its subdivisions, Polynomial Texture Mapping (PTM), developed in 2001 at Hewlett Packard Laboratories (Malzbender, Gelb and Wolters 2001). The technique captures a series of raking and oblique light images of a static object with a static digital camera at constant exposure, either by the dome method or the Highlight-based method. After processing using specialised software (RTI builder), it represents the 3D reflectance properties of objects per pixel. Viewers, such as the RTI viewer (ISTI-CNR/CHI RTI Viewer) and the PTM viewer (HP Labs PTM Viewer), enable interactive manipulation of the lighting position and enhancement of the final image through different rendering modes. Via RTI, experts and the public can experience views of the artefacts in a new way that is superior to visual examination or handling under a microscope lens.

2. Previous Work

The technique significantly contributes to analysis, conservation and representation
There are many interesting applications in the field of cultural heritage, based on its ability to acquire and represent the 3D reflectance properties of objects. Related projects are the Antikythera Mechanism (Malzbender and Gelb 2009), the RTISAD Project of the University of Southampton and the University of Oxford (Earl et al. 2011), Reflectance Transformation Imaging in Canada at Queen’s University (RTiiCAN) (Gabov and Bevan 2011), and Reflectance Transformation Imagining (RTI) in art conservation, a collaboration between the Fine Arts Museums of San Francisco and Cultural Heritage Imaging (CHI). Previous work includes the application of RTI in paintings, scrolls, wood, metals, ethnographic materials, sculpture etc (Caine and Magen 2011; Padfield, Saunders, and Malzbender 2005; Karsten and Earl 2010). Microscopic RTI, using the highlight or the dome method, gave interesting results and this lead to its broader application in the cultural heritage sector, particularly in conservation practice (Kotoula and Earl 2012).

Although the RTI community is increasing in number, as well as expertise, applications of RTI in representations before and after conservation operations are limited.

3. Methodology: Microscopic RTI

To fully appreciate the application of RTI in conservation practice, one should take into consideration conservation needs for limited human-object interaction, high quality and affordable visual analysis, microscopic levels of detail and advanced documentation. There is no doubt that methodological developments are necessary in order to meet those needs. One of the most obvious of these, even from the earlier stages of our research, was the microscopic level of detail required for conservation. Microscopic lenses are considered as an extension of the conservator’s eyes. Initially, the Highlight RTI method was tested under the microscope, which proved to be extremely difficult, due to the limited space available for setting up the scene, and considering the presence of the sphere, along with its shadows. Another disadvantage is the microscope arm, which excludes almost 25% of the lighting positions.

Based on the above observations, an experimental microscopic RTI dome was built. The first one was a simple wooden structure, equipped with 18 LED lights (Kotoula and Earl 2012). Following promising results, a more complex metal dome, suitable for microscopic data capture, was designed and built, capable of capturing 64 microscopic images, four images from different angles to the object’s plane, from each of the 16 positions of the periphery of the dome. In addition, in order to enhance the captured data, the four lights were not placed in vertical order. The dome can be turned 90 degrees, in order to place the artefact securely on a stable plate. A metallic ring, joined with the dome, is located over the artefact plate. The
metallic dome, equipped with four LED lights, is moving clockwise around the artefact’s plate. It is impossible for the moving dome and the ring to be separated or to touch the artefact (Fig. 1). Apart from microscopic RTI, the same dome can also be used for macroscopic RTI data capture, for example for coins or small fragments.

4. Results and Discussion

4.1 An Overview of RTI in Archaeological Materials

The results of our experimental study, including macroscopic and microscopic RTI, making use either of the dome or the highlight method, lead to better understanding and more advanced documentation of crucial features for the study of the artefacts’, finds analysis and conservation. The enhanced visual analysis enables identification of materials based on their optical properties, and it can highlight points on the surface, which should be further studied with physicochemical analytical techniques. Also, it enables advanced morphological analysis of patinas, depositions and encrustations, significantly assisting in physical damage characterization. Hidden or low relief decorative elements, such as inscriptions, engravings and impressions can be studied and documented easily. In addition, RTI can emphasize and underline material evidence, which can influence the artefacts’ interpretation and characterization. In this way, it can help test hypotheses, and be used as a tool for conservation decision making. Each feature on the surface of artefacts can be documented and examined, so as to reach conclusions for every episode of its four stage biography; 1) previous treatments or “museum life”, 2) use and wear, 3) decay and 4) working, manufacture or fabrication evidence (Caple 2000).

The initial experimentation is aimed at investigating RTI visualization, using a variety of material types in a wide range of conservation and preservation states, in comparison to traditional raking light imaging. Artefacts derived from the Hellenistic-Classical Derveni cemetery in Macedonia, Greece, currently located in the Archaeological Museum of Thessaloniki, which were either exhibited or stored under controlled environmental conditions, were used as a case study. Raking light is an imaging technique widely used in conservation examination. The results indicated that the disadvantages of raking light cannot be overlooked. Raking light examination should not be considered in the case of vulnerable and fragile artefacts, because the handling required for a complete examination contravenes basic conservation rules and codes of ethics, as it may cause damage. Although raking light imaging is preferable from a preventive conservation point of view, due to the limited human-object interaction, the results are rather disappointing, because of the large number of images needed for a complete examination, and the problems caused by the manipulation and examination of the captured images. No single captured image can demonstrate the characteristic features of archaeological objects, such as three-dimensionality and geometric complexity, in addition to the surface topography of deteriorated material.

On the contrary, the study highlights the advantages of RTI. When applied to ceramics, RTI visualization enabled advanced documentation and examination of previous treatments, gap-fillings, crumbling ceramic material, rough surface finish with scratches and pits, craquelure, flaking, loss and cracking of gold and/or colour decoration, salt efflorescence and depositions, as well as low relief decoration features and engravings (Fig. 2). For glass and stone vessels RTI emphasized manufacture details. In addition, usual degradation effects and features become apparent, such as salt efflorescence, cracking, flaking, pitting and losses and inscriptions (Fig. 3). With bronzes, the focus is the examination of patinas, encrustations and depositions,
Figure 2. Ceramic roundel, burned, cracked with losses and minor remnants of gold decoration, Derveni, tomb A (diam. 3cm). Clockwise from top left, digital image, RTI visualization and microscopic RTI visualization of 1.5 mm.

Figure 3. Glass vessel, alabastron, of dark blue glass, decorated with dense feather pattern in white, blue and yellow, Derveni, tomb D (length 8.9 cm). Digital image (left) and RTI visualization of details (right).

Figure 4. Bronze mirror, fragmented and highly corroded, with incised decoration, Derveni, tomb Z (length 6.6 cm). Digital image (top) and RTI visualization of details (below).

Figure 5. Fragment of a wooden pyxis cup, Derveni, tomb B (1cm across). Digital image (top) and RTI visualization of details (below).

Figure 6. Leather fragment attached to a copper alloy ring, Derveni, tomb B (length 4cm). Digital image (left) and RTI visualization of details (right).
crevices and warts via RTI (Fig. 4). With organic material, the most common effects noted were deformations and physical damage, including shrinkage, warping, cracking and losses. In the case of wood, the observation of the grain direction is considered important, as it can lead to identification of the wood type, while low relief details in bone can assist osteoarchaeological studies (Fig. 5). It is also worth mentioning the delamination and mineralization effects observed on the RTI visualised leather samples (Fig. 6).

4.2 Identification, Examination, Cleaning Monitoring and Documentation of Coins

RTI is considered to be a powerful documentation tool for cultural heritage applications. Previous work in the field of RTI in numismatic studies demonstrated advantages in identification and documentation issues in comparison to traditional approaches, as well as improved communication and dissemination abilities (Mudge et al. 2005; Gabov and Bevan 2011). In a recent study, twenty five Roman or ancient Greek copper alloy and silver coins were visualised in .ptm and/or .rti format before, during and after mechanical cleaning under the microscope, in an attempt to test RTI’s efficacy in conservation documentation and monitoring of remedial conservation operations. A numismatic collection was chosen for this study because of the high demands of coins documentation and the inefficiency of traditional methodologies. A further reason was the thorough details coins include, which can have a tremendous archaeological impact, for example in assisting dating. The results indicated that RTI not only assists coin identification, but also enables advanced conservation documentation and monitoring of cleaning operations. Moreover, RTI’s use in the field of preventive conservation, as well as its significance for providing access to cultural heritage collections is advantageous, in particular in the case of coins, given problems of small size and low relief detail.
RTI’s ability to emphasize surface variation proved to be extremely important for coin identification and dating. A characteristic example is the lettering of a silver diobol of the Macedonian Kingdom, Amyntas III, c. 393 - 369 B.C (Fig. 7). In addition, RTI offers an advanced method for the examination and condition reporting of surface effects, such as corroded, pitted surfaces with encrustations and scratches. A copper alloy coin, from the Thracian city of Maroneia, represents Asclepius standing, head left, leaning on a serpent-entwined staff. RTI enhances legibility and condition reporting (Fig. 8). Other effects due to conservation treatment can be captured using RTI. A silver Roman imperatorial denarius of Julius Caesar, minted in North Africa, 47–46 BC, was chemically cleaned. As a result, it presents a stripped, porous, rough surface. Moreover, the scene represented on the reverse, Aeneas advancing to front, holding Palladium in the palm of his right hand and carrying his father Anchises on his left shoulder, as well as the legend “CAESAR”, is hardly legible. The legend and the scene appear more vivid in the RTI snapshots than in digital images (Fig. 9).

RTI’s is also useful for recording and visualization of the coin’s “museum life”, which is an inseparable part of its biography, consisting of the three basic stages before, after and during conservation. Hence, monitoring of the conservation process as well as the condition of the coin is possible. The goal of numismatic conservation is to remove the active corrosion products, which can cause further degradation, even under controlled environmental conditions, in addition to revealing the representations and legends. This helps safeguard material integrity, and to re-establish aesthetic values, so that the exhaustive examination of the coins’ surface after conservation is imperative. RTI can be considered as the most effective way of evaluating the success of such a treatment. The autonomous coin of the Koinon of Macedon, c. 3rd century AD., which represents the head of Alexander III on the obverse, and a horseman charging right, was visualised before and after treatment (Fig. 10).

4.3 RTI and Preventive Conservation

Due to the lack of preventive conservation measures, artefacts may rapidly age and degrade. Apart from the control of environmental conditions, mishandling is one of the factors of long-term deterioration, responsible for damage after excavation. In general, less handling of artefacts is recommended, not only because of the risk of damage, but also because contact with human skin can encourage corrosion due to fatty acids. As well, during close up examination, the humidity from human breath can cause alterations and disturb the equilibrium of relative humidity and temperature.

Where the protection and the stability of artefacts are concerned, RTI visualization, as well as photography-based digital imaging techniques or lab recording strategies, can be considered as techniques that respect the material culture of the past. In particular, RTI can play a crucial role as a preventive
conservation measure, not only because it significantly improves the accuracy of digital representation, but also because it delivers archaeological information better than the artefact alone.

4.4 RTI and Practicality

A complete critical analysis of the application of RTI in conservation must consider practicality, which along with accessibility, durability and integrity forms the major conservation objectives (Watson 2010). The practical implementation of RTI must take in to account economic considerations. It is affordable, as only standard photographic equipment is required for data acquisition, and the necessary software for processing is available online, and can be downloaded without cost for non-commercial. No extensive training is required. Expertise can be acquired through videos and detailed guides to highlight image capture. Significantly the software involved in processing and viewing reflectance transformation data has a friendly user interface.

5. Conclusions and Future Work

The role of RTI in non-destructive examination, analysis, conservation documentation and monitoring, as well as in preventive conservation, has been discussed, using artefacts derived from Macedonia as a case study, dated back from ancient to Roman times. The conclusion is that RTI fulfils the objectives of a conservation project. RTI cannot be considered as an unnecessary ‘modernisation’ of processes, dictated by technological developments, but is actually a useful tool. Hence, the application of RTI disproves critical views within the conservation literature which argue against the role of science and technology, and suggest an unbalanced view of science and technology in the field of practical conservation.

Microscopic RTI, which meets conservation needs and furthermore signals interesting developments in data capture methodology, was highlighted. The synergy of RTI and microscopy can broaden the applications of the former in the cultural heritage sector, particularly in lab recording for conservation purposes. RTI’s ability to help narrate the object’s “museum life”, demonstrated via the visualizations of the coins in different stages of a conservation project, considerably enhances conservation documentation. The development of a methodology for automatic comparison of pre- and post-treatment RTIs would provide a powerful alternative solution to documentation and condition monitoring.

RTI is currently under development in various research directions, including the enhancement of final images, which improves the perception of details, features, and overall shape characteristics (Palma et al. 2010). As well multi-view RTIs, promote a complete approach towards documentation of objects’ provenance information, and are more scientifically reliable, but also assist reprocessing, and improvement of the annotation tool, which facilitates more effective scientific cooperation (Mudge et
The application of multispectral RTI can potentially provide information for the surface variation of elements seen under ultraviolet or infrared radiation, enabling advanced examination, as well as assisting interpretation. Infrared reflected and transmitted imaging provides information about layering of painted artefacts and works of art, epigraphy on textile, wood, stone, metals and ceramics. It penetrates aged and/or oxidised varnishes and colour layers. It reveals the stages of creation of an artefact, and details hidden beneath coloured varnishes or degradation effects, such as encrustations. It is extremely helpful for the detection of “pentimenti”, under-drawings and over-paintings (older or modern). In addition, it is possible to identify and differentiate materials. It is used in authenticity studies and condition reports. Materials with similar optical characteristics cannot be differentiated with the naked eye. Ultraviolet photography provides an easy way to characterize materials and examine surfaces. Not only the presence of surface coating, over-paintings and other modern additions can be determined, but also the conservation state of these can be examined and recorded. Flaking loses and abrasions of an invisible transparent varnish, modern repairs, and techniques of manufacture/application of materials can also be easily investigated (Warda et al. 2011). The advantages of multispectral imaging, in synergy with RTI, can potentially enhance multispectral photographic examination, offering the possibility to visualize not solely the hidden inner or outer layers, but also its three-dimensional reflectance properties. The most important research problem is the unevenness in illumination, which can appear as different absorbencies. The question is whether multispectral RTI can follow the same data capture methodology.

Acknowledgements

The author gratefully acknowledge the Archaeological Museums of Amphipolis and Thessaloniki (Hellenic Republic-Ministry of Culture and Sports) and the Archaeological Computing Research Group (University of Southampton). The author would like to express many thanks to Dr. Graeme Earl, Dr Despina Ignatiadou, Aikaterini Peristeri and Maria Kyranoudi for their support towards the completion of this study.

References


Automatically Recognizing the Legends of Ancient Roman Republican Coins

Albert Kavelar, Sebastian Zambanini and Martin Kampel
Vienna University of Technology, Austria

Abstract:
This paper presents an approach towards an automatic recognition of ancient Roman Republican coin legends. Standard Optical Character Recognition (OCR) techniques depend on horizontal text alignment and binarization for text segmentation and thus fail to recognize ancient coin legends due to their arbitrary orientation and the complex shading effects on the highly specular coin surfaces. Hence, the proposed method is rooted in object recognition and employs local image features for character description. SIFT descriptors for densely sampled locations on coin images are computed and classified using Support Vector Machines (SVM) individually trained for every character class. The resulting scores are combined into words, using the pictorial structures approach, and the most probable word of a finite lexicon is chosen. The word recognition rates achieved in experiments on a set of 180 images showing Roman Republican coins range from 29% to 53%, depending on the lexicon size used.

Keywords:
Computer Vision, Ancient Coins, Coin Legend Recognition

1. Introduction

Numismatics is the scientific discipline dealing with all forms of currency and other payment media as well as their historical aspects. Besides studying the distribution of coin hoard finds or the identification of coins, coin classification plays a crucial role. By classification, numismatists refer to the association of a specific coin to a certain number according to reference books such as the Crawford catalogue (Crawford 1974; Kampel and Zaharieva 2008). This is a complex task, which requires years of experience in the field of numismatics (Kampel and Zaharieva 2008). Apart from material analysis, classification primarily relies on the inspection of images and textual inscriptions on the coin. Hence, coins can be classified based on digital photographs to which computer vision methods can be applied, thereby assisting numismatists in their work. In order to assign an ancient coin to a certain class, numismatists have to identify and discern their characteristics, highlighted in Fig. 1. One of the most significant features used for classification is the textual inscription called “legend” which, for the Roman Republican coins considered in this work, often reads the name of the issuing authority. Consequently, a legend recognition technique could substantially contribute to a fully-fledged coin recognition system. Moreover, an automatic classification system could be integrated in online platforms for hobby collectors allowing them to classify their coins and gain background information on their collection by uploading a digital photograph showing the coin. Furthermore, it could help to stem the illicit trade of stolen coins on online auction platforms by frequently checking the respective websites for uploaded coin images and comparing them to a database of stolen coins, a problem previously addressed by the COINS project. (Arandjelovic 2010; Zaharieva, Kampel and Zambanini 2007b).

This work focuses on Roman Republican coinage but the proposed technique could easily be applied to other coins such as Roman Imperial coins or modern coins, as long as the used classifier is trained for the letters that occur in the respective legends. We decided to focus on Roman Republican coins because Crawford (1974) provides a standard-reference work giving a comprehensive overview of the different coin types of this period. Moreover, the presented research is part of the ILAC
project (Zambanini, Kampel and Vondrovec, to appear) carried out in collaboration with the Kunsthistorisches Museum Wien (Vienna Museum of Fine Arts) which has got one of the largest collections of Roman Republican coins worldwide.

Fully automated classification systems have been developed for modern coins, e.g. Dagobert (Nölle et al. 2003) or Coin-O-Matic (van der Maaten and Poon 2006). However, up until now, there is no existing classification system for ancient Roman Republican coins. Due to their manufacturing process, ancient coins differ substantially from their modern counterparts and hence pose different challenges to computer vision. Ancient coins were cast or struck from handcrafted dies showing the inverse of the image to be minted on the raw piece of metal called flan, as illustrated in Fig. 2. This manufacturing process results in coins having individual minting marks, different shapes, off-centred or cropped imagery and misaligned obverse and reverse images. Hence, each coin is unique and the intra-class variability is high which impedes the application of methods for modern coins (Nölle et al. 2003; van der Maaten and Poon 2006), as they assume uniform background and identical appearance for coins of the same class.

At first glance one might wonder why we do not address the problem with traditional OCR methods, since they are known to perform very well on scanned text (Wang and Belongie 2010). However, conventional OCR methods for scanned documents as presented by Plamondon or Vinciarelli do not account for the challenging conditions of coin legends (Plamondon and Srihari 2000; Vinciarelli 2002). These techniques depend on horizontally aligned text, an assumption not met by coin legends as they can be oriented arbitrarily (see Fig. 3). Moreover, traditional OCR methods rely on binarization and thus fail on coin images, since the only visible information is intensity changes in the image resulting from sharp surface normal changes in the coin relief (Arandjelovic 2010). Hence, text and background colour are identical, leading to segmentation errors in binarization and subsequent misclassifications. Therefore, the presented approach employs a binarization-free character recognition technique based on local image features. To increase word recognition rates, a dictionary comprising the finite set of possible legend words is incorporated in the recognition process. The presented technique is similar to the scene text recognition method introduced by Wang et al. (2011).
The remainder of the paper is structured as follows: In Section 2 related research regarding character and word recognition in natural scenes is reviewed. In Section 3 our proposed legend recognition system is presented. Section 4 summarizes the character and word recognition experiments carried out; and finally, Section 5 concludes the paper with an outlook on the next steps in our research.

2. Related Work

There are several approaches towards image-based classification for ancient coins (Arandjelovic 2010; Huber-Mörk et al. 2011; Zaharieva, Kampel and Zambanini 2007a). However, these approaches do not exclusively consider the legend but rather use the entire coin in the classification process. To our knowledge, there have not yet been attempts to recognize ancient coin legends. As traditional OCR methods are inappropriate, our review concentrates on the related field of scene text recognition which aims at detecting text in unconstrained images of natural scenes (Wang and Belongie 2010). De Campos et al. (2009) mention varying font style and thickness, different foreground and background colour combinations, lens distortion, illumination changes and arbitrary image resolution and compression as the challenging conditions encountered in natural scene images. These conditions need not be considered when dealing with scanned document images. According to them, a scene text recognition system comprises the following steps: (1) text localization, (2) character and word segmentation and (3) integration of language models and context. However, their method solely addresses character recognition and assumes cropped images of English and Kannada characters. In their experiments with different local image features and classifiers, they come to the conclusion that combining all features and using Multiple Kernel Learning (MKL) yields the best results, closely followed by geometric blur (Berg, Berg and Malik 2005) combined with 1-against-all SVM tests. Diem and Sablatnig (Diem and Sablatnig 2009) introduced a method for the recognition of Glagolitic characters in degraded documents using different local image features. Every keypoint is classified employing a one-against-all SVM test for keypoint classification. After clustering keypoints, a voting scheme assigns each character to its most probable class. They carried out experiments on a dataset of images showing glagolitic manuscripts and evaluated the character recognition performance of several rotationally invariant local image descriptors, namely SIFT, SURF, GLOH, PCA-SIFT and gradient moments. The best classification rate was achieved using SIFT features and lies at 38.8%.

Wang and Belongie present an approach towards scene text recognition operating on cropped word images (Wang and Belongie 2010). They use Histogram of Oriented Gradient (HOG) features (Dalal and Triggs 2005) with a Nearest Neighbour (NN) classifier. Character
segmentation is performed by shifting all templates of every class across the image and computing the normalized cross correlation between the template’s HOG features and the ones at the window location. The maximum value for every class is used as score at this location. The candidate locations remaining after a non-maximum suppression (NMS) are then combined to words using Pictorial Structures (Felzenszwalb and Huttenlocher 2005). Their approach was extended to a fully-fledged scene text recognition system by Wang et al. (Wang, Babenko and Belongie 2011). In contrast to their previous work, they operate on entire scene text images rather than on cropped word images and use Random Ferns instead of an NN classifier. Their experiments show that the presented method outperforms commercial state-of-the-art OCR systems and that the omission of the text localization step does not significantly affect the word recognition performance.

### 3. Methodology

Text localization methods for natural scene images aim at detecting machine-printed text in images, such as the approach introduced by Clark and Mirmehdi (Clark and Mirmehdi 2002). Wang and Belongie focus on road signs or words printed on product labels in grocery stores (Wang and Belongie 2010; Wang, Babenko and Belongie 2011). In these scenarios, the text has a colour contrasting with a uniform background, an assumption not met by coin legends. Therefore, contrary to the recommendation of de Campos et al. (2009), the proposed legend recognition pipeline depicted in Fig. 4(a) does not foresee a text localization step. However, as Wang et al. (2011) showed, omitting the text localization step has no significant effect on word recognition rates.

#### 3.1 Training

As we use an offline machine-learning approach, an initial training step is required (see Fig. 4(b)). Since characters on ancient coins can be oriented arbitrarily, it is important to use a local image descriptor providing rotational invariance. We use SIFT features, as they provide rotational invariance and have previously been used successfully for character recognition (Diem and Sablatnig 2009). Due to the wide range of different illumination conditions and directions and the various degrees of wear, SIFT keypoint localization is not robust for our training set of cropped characters. Therefore, only one SIFT descriptor scaled to cover the entire image for a centred keypoint is computed, as illustrated in Fig. 5. Considering only the uppercase letters encountered in Roman Republican

---

*Figure 4. (a) Legend Recognition Pipeline (b) Training of SVMs.*
coin legends gives us an overall of 18 different character classes. 50 images per class are fed to the machine learning process, as illustrated in the top of Fig. 4(b). For each of these images, a single SIFT descriptor is computed and used to train an individual SVM for every class. The 18 readily trained SVMs are stored in a database to be used in later character recognition (depicted at the bottom of Fig. 4(b)). Two different SVM databases are trained: one using SIFT features oriented according to the dominant gradient direction of the underlying image patch, hence providing rotational invariance and one using fixed orientations that lead to better character recognition results in the re-scoring step (see Section 3.2.3).

3.2 Legend Recognition

The proposed legend recognition architecture is illustrated in Fig. 4(a); each of its three major steps depicted is explained in detail in the following subsections. As in the scenario presented by Wang et al. (Wang, Babenko and Belongie 2011), we face a limited set of known words to be recognized. Thus, we can pass a lexicon comprising all legend words to the recognition pipeline for an increased recognition performance. Initially, input images are scaled to a fixed size of 384×384 pixels for reduced computation times (Step i. in Fig. 4(a)). Moreover, SIFT descriptors only have to be computed at a single scale, since using a fixed image size leads to a rather constant font size for all test images.

Keypoint Extraction

The resized image is passed on to the keypoint extraction step (Step 1 in Fig. 4(a)), where areas containing no relevant information (i.e., homogeneous regions without edges) are eliminated by the application of an entropy filter, as they do not contain legend information. Thresholding the filtered image results in a binary map marking Regions of Interest (ROI) to be considered in the subsequent steps. The hereby found ROIs are densely sampled by taking every second pixel in the x- and y-direction. These sample points represent character candidate locations.

Keypoint Classification

For every character candidate location passed to the keypoint classification step (Step 2 in Fig. 4(a)), a SIFT descriptor with dynamic orientation is computed and tested against all SVMs trained accordingly. Consequently, each candidate location is now associated with 18 different scores indicating how likely it is to encounter a character of the respective class.

Word Detection

Finally, the word detection step (Step 3 in Fig. 4(a)) searches the downsampled image for the words given in the lexicon. This is done by combining the scores of the character candidate locations to lexicon words. Like Wang et al. (Wang, Babenko and Belongie 2011), we employ pictorial structures (Felzenszwalb and Huttenlocher 2005) for finding the optimal configuration of a certain word. As described by Wang and Belongie (Wang and Belongie 2010), pictorial structures are like a mass-spring model which can be used for object recognition. The basic idea is to describe an object by its
sub-parts and to match them to image locations whilst minimizing the costs associated with the placement. In the case of coin legends, we consider words as objects and characters the parts a word is built of. Characters can be placed at the character candidate locations found in the keypoint extraction step. Thus, when scanning an image for a specific lexicon word, we try to find the least expensive way to place its letters in the image. There are two kinds of placement costs. First, the costs resulting from placing a character at a certain candidate location, i.e., the score computed for this class and location in the keypoint classification step. Second, the Euclidean distance in pixels between two consecutive letters. Obviously, there are upper and lower bounds to the distance as letters are not allowed to be overlapping and too large a distance implies that two characters cannot belong to the same word. These bounds were determined empirically. More formally, the problem can be formulated as a graph minimization problem. Let \( G = (C, E) \) be the directed acyclic graph representing the configuration of a certain n-characters long word \( W \) in an image where \( C = \{c_1, \ldots, c_n\} \) is the set of characters and \( E = \{e_1, \ldots, e_{n-1}\} \) denotes the set of edges connecting two consecutive characters. Let \( L = \{l_1, \ldots, l_n\} \) represent a certain configuration found within the image, i.e., the \( i \)-th character of the word \( W \) is placed at a certain image location \( l(x_i|y_i) \). The placement costs equal the score computed for the character \( c_i \) at the location \( l_i \) in the keypoint classification step. The weights of the edges \( e \in E \) are defined by the Euclidean distance \( d(c_j, c_{j+1}) \). Among all possible configurations \( L \), the optimal solution \( L_{opt} \) having the lowest costs is found using dynamic programming, as suggested by Wang, Babenko and Belongie (Wang, Babenko and Belongie 2011).

However, resulting from the rotational invariance of the SIFT features used for describing the characters, a found word configuration may have inconsistent character orientations, as shown in Fig. 6. Therefore, subsequent re-scoring is applied to the best 50 configurations found for a certain word. In this step, the known edge directions of the found word configuration \( L_k \) are used to re-compute SIFT descriptors fixed to the according edge direction at each location \( l_{ki} \). That is, the SIFT descriptor at the location \( l_{ki} \) is tested against the SVM for the character \( c_i \) with fixed SIFT orientations. If one character orientation in the original configuration \( L_k \) should not match the according edge direction, it will receive a very low score. Consequently, words with inconsistent character orientations are penalized and thus rejected. To allow a comparison of the scores achieved for words of different lengths, they are eventually divided by the number of characters for the respective word. A threshold applied to the final word scores helps filtering out randomly found nonsense words. Since at the current state of our research only one word is sought in an image, it is assigned to the lexicon word which causes the lowest costs.

4. Experiments

In this sections we present results for isolated character classification and for legend recognition. Our testdata comprises cropped images of ancient Roman Republican coins.

4.1 Character Classification

In this experiment, we test the appropriateness of the way we employ SIFT descriptors for character representation. All images used are scaled to the uniform size of 100×100 pixels and show a single character.
In some cases, parts of adjacent characters are visible due to the letter’s aspect ratio (see Fig. 7(a)). As described in Section 3.1, two SVM databases are required, one with SIFT descriptors oriented at a fixed orientation of 0° and one where the orientation of the descriptor is based on the dominant gradient orientation of the image (and thus providing rotational invariance). Hence, we experiment with fixed and dynamic SIFT orientation settings. Moreover, we try to compensate for the artefacts introduced by different lighting directions (see Fig. 7(b)) by accumulating the SIFT descriptor values for opposing directions (e.g., the values for 45° and 225° are summed up) resulting in a bisected descriptor length which we will refer to as SIFT180 as opposed to SIFT360 denoting the unmodified descriptor. Additionally, the size k of the training set per character was varied in our benchmarks (sizes of 15, 30 and 50 were used). The test set used in the character recognition benchmarks comprises 80 images, each belonging to one of 18 character classes. The results achieved are listed in Table 1. Generally speaking, the recognition rate increases with the size of the training set. Moreover, using a fixed orientation significantly increases the recognition rates. This results from the fact that the dominant gradient orientation heavily depends on the lighting direction, as depicted in Fig. 7(c), which illustrates how the computed SIFT descriptor orientation follows the lighting direction for certain characters. This can partly be compensated with SIFT180, as the orientation’s degrees of freedom are reduced; a fact reflected in the increased recognition performance for SIFT180 compared to SIFT360. Consequently, the best result lies at 72% and is achieved with SIFT180 using a training set size k=50.

4.2 Legend Recognition

Word detection was benchmarked on a test set of 180 images of pre-segmented ancient Roman Republican coins, i.e., no additional clutter, such as rulers for size reference, is visible in the image. The images were tested with lexica of different sizes comprising a certain number of randomly added distractor words chosen from a list of 35 possible words and the word actually depicted in the image. Lexica of 5, 10, 20 and 30 distractor words were used. The results achieved for the presented legend recognition pipeline are listed in Table 2. Best match means that the word with the lowest costs found for an image matches the correct word. Within Top-3 indicates that the

<table>
<thead>
<tr>
<th>k=15</th>
<th>k=30</th>
<th>k=50</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIFT360, fixed orientation 45% 49% 68%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIFT180, fixed orientation  42% 64% 72%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIFT360, dynamic orientation 31% 33% 42%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIFT180, dynamic orientation 25% 33% 34%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Character Classification Results.

<table>
<thead>
<tr>
<th>n=5</th>
<th>n=10</th>
<th>n=20</th>
<th>n=30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best match, initially 31.7% 21.7% 17.8% 13.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Top-3, initially 62.2% 39.4% 21.7% 18.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best match, after re-scoring 53.3% 42.8% 34.4% 28.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Top-3, after re-scoring 81.8% 66.1% 57.8% 51.7%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Legend Recognition Results.
configuration found for the actually depicted word is among the three words causing the lowest costs. Obviously, the recognition rates drop as the lexicon size increases because the larger the lexicon, the less likely it is to find the right word by chance. However, the results show that re-scoring improves the recognition performance significantly, since it eliminates many false positives.

5. Conclusions

Our preliminary results show that this work is a first step towards an automated legend recognition system for ancient Roman Republican coins. We employ object recognition techniques for character recognition, and achieve promising classification rates (72% for 18 classes) in comparison to the work of de Campos et al. (55% for 62 classes on their Natural Images dataset) (de Campos, Babu and Varma 2009) and Wang et al. (64% for 62 classes on the ICDAR dataset) (Wang, Babenko and Belongie 2011). Nevertheless, the word recognition performance still needs to be improved in order to contribute to a classification system for ancient coins. In the future we will experiment with other local image features suggested in related research and thereby hope to reduce false positives in the character recognition step. An increased character classification will also lead to a better overall legend recognition. Moreover, we will classify characters using Random Ferns for a reduced computation time since they offer real multi-class support.

Acknowledgements

This research has been supported by the Austrian Science Fund (FWF) under the grant TRP-140-N23-2010 (ILAC).

References


Multispectral Imaging of Historic Handwritings

Fabian Hollaus
Vienna University of Technology, Austria

Abstract:
Multispectral imaging has proven its usefulness for the examination of ancient manuscripts, since it enhances the legibility of vanished or erased writings. Hence, this non-invasive conservation technique facilitates the work of philologists. This paper is concerned with the acquisition and digital processing of multispectral images containing historic writings. The manuscript pages examined contain an overwritten historic text and an overlying handwriting, which is considerably younger than the overlying text. The younger texts are visible under all wavelengths utilized, while the older texts are best legible under UltraViolet illumination. This work presents efforts, which have been taken, in order to make the ancient writings readable. At first the image acquisition setup is detailed and the image processing methods, which have been applied, are explained in the second part of the document.

Keywords:
Multispectral Imaging, Inpainting, Palimpsest, Document Analysis, Computational Photography

1. Introduction

This work deals with the automated recovery of palimpsests. Palimpsests are ancient writings on parchment, which have been overwritten, since parchment was an expensive material in former times. The folios, which are used in this work, were additionally erased before they were overwritten. Due to this erasure, the underwritings are hardly legible under tungsten illumination (Easton et al. 2003) have shown that the underwritings of the Archimedes palimpsest are most visible under UltraViolet (UV) illumination, while they are barely visible under red light. Furthermore, they developed a pseudocolor technique, which visualizes the ancient and the younger text.

An example for this pseudocolor technique is given in Figure 1. The pseudocolor approach has been applied on a palimpsest, which has been imaged by our project team in the National Library of Sofia, Bulgaria. The shelfmark of the manuscript is NBKM880. The pseudocolor image has been generated by putting the red channel of the white light image shown in Figure 1 (a) into the red channel of the pseudocolor image, given in Figure 1 (c). The remaining channels of the pseudocolor image are filled with the blue channel of the UV fluorescence image shown in Figure 1 (b).

This example shows that multispectral imaging enhances the visibility of palimpsest underwritings – as it is also noted in (Easton et al. 2003) and (Fischer and Kakoulli 2006). However, the legibility of such texts is nevertheless constrained, since they are partially occluded by younger writings. This work aims (among others) at closing the stroke gaps stemming from occlusions automatically and thus simplifying a text analysis by scholars.

The recovery system introduced consists of two consecutive steps: At first the occluded regions are found and their locations are encoded in an inpainting mask. Afterwards the affected regions are recovered by applying a statistical inpainting approach. The detection of the occlusions is fulfilled with the help of a binarization algorithm (Su et al. 2010), which is suited for the segmentation of degraded and historical texts. The algorithm is applied on red light images, since such images exhibit barely the underwritings.

The ensuing inpainting step is instead fulfilled on photographs of parchments, which have been exposed to UV light. Digital inpainting
defines the automated filling of an unknown or damaged region, which is given in the form of a binary mask. Application fields for inpainting are for instance image retouching (Bertalmío et al. 2000) or object removal (Komodakis and Tziritas 2007). Our system makes use of an Markov Random Field (MRF) model, which captures the statistics of handwritings. The MRF modelling approach chosen is called Fields of Experts (FoE) framework and was proposed by Roth and Black (Roth and Black 2005). Contrary to previous MRF modelling approaches, the FoE framework allows for an offline training of high-order MRFs.

This paper is organized as follows. In the following section our multispectral imaging (MSI) is detailed. Afterwards the binarization algorithm and the inpainting technique are introduced. Section 5 provides an evaluation of the palimpsest inpainting technique and Section 6 contains concluding remarks and an outlook on future work.

2. Multispectral Imaging System

Our acquisition setup is illustrated in Figure 2. The images are captured by two different cameras: The camera on the left in Figure 2 is a grayscale near infrared camera with a spatial resolution of 4000 x 2672px. The spectral response range lies between 300 nm and 1000 nm. The camera on the right is a single-lens reflex Nikon D2Xs, with a resolution of 4288 x 2672px. The Nikon camera is used for white light and UV fluorescence photographs and hence a filter is used in order to filter out the reflected UV light.

The acquisition setup evolved during the course of time. In the first project phase the multispectral scan was acquired with optical filters that are build in a filter wheel, which is mounted in front of the Hamamatsu camera. Contrary, in the current project phase we make use of two Eureka!Light™ LED panels (Equipoise imaging, Archimedes project (Easton, Knox and Christens-Barry 2003)). These panels enable an acquisition in 13 narrow spectral ranges, which are shown in Figure 3. Since the LED panels provide filtered light, there is no further need to filter the reflected light and hence the photographs are captured without a filter, except for two special imaging kinds – namely UV reflectography and UV fluorescence. For images of these particular kinds it is still necessary to filter the reflected respectively the fluorescent UV light. Compared to the previously used tungsten illumination, the LED panels have the advantage that the heat put on the manuscripts is decreased.
The manuscripts are placed on a plate, which is affixed on a linear unit. The linear unit is used in order to move the books between both cameras. The usage of the linear unit is convenient for the imaging of bounded books, since the manuscripts do not have to be moved manually, but just the pages have to be flipped.

In Figure 4, a portion of the above mentioned Bulgarian palimpsest imaged with our MSI system is given. Similarly to (Easton et al. 2003) we noticed that the ancient writing is most visible in UV fluorescence images. It is also apparent that the overwritten text disappears faster than the younger text with increasing wavelengths. This can be attributed to the fact that the ancient writing reflects more red light than the younger writing.

3. Inpainting of Palimpsest Writings

The palimpsest recovery described in the following has been evaluated on images that have been captured in the Archimedes palimpsest project (Easton et al. 2003), since philologists have generously provided a public available transcription of the manuscripts (Netz et al. 2004). This enabled us to create ground-truth data of the overwritten text. The recovery technique has not been applied on the Bulgarian manuscript yet, but we are planning to evaluate the palimpsest recovery technique on this writing either.

3.1 Generation of the Inpainting Mask

Su et al. (Su et al. 2010) have recently proposed a segmentation technique for ancient text documents. The algorithm considers the image contrast, which is defined by the intensity maximas and minimas within local neighborhoods. These local intensity extrema are used for the generation of a contrast image, which is defined as:

\[
D(x, y) = \frac{f_{\text{max}}(x, y) - f_{\text{min}}(x, y)}{f_{\text{max}}(x, y) + f_{\text{min}}(x, y) + \epsilon}
\]

where and are the maximum and minimum intensity values in a local neighborhood. The denominator consists of an infinitely small number , which prevents a division by zero, and a normalization term . This normalization term is used in order to lower the influence of background variations and to enable a detection of faded out characters. In Figure 5 (a) a parchment portion is shown, along with the corresponding contrast image in Figure 5 (b). The algorithm proceeds with an identification of the stroke boundaries. It is assumed that the local contrast is especially high at stroke...
boundaries - compared to the background - and the character boundaries are detected by applying a global Otsu (Otsu 1979) threshold. The foreground pixels found at this juncture are called high contrast pixels. The overall segmentation is concluded by a detection of the foreground pixels, which are in the near of the high contrast pixels. A pixel is classified as a text pixel if the following conditions are met: On the one hand the pixel must have at least a certain number of high contrast pixels in its local neighborhood. On the other hand the intensity value of the pixel must be equal or smaller than the average intensity value of the high contrast pixels within the same neighborhood window. Figure 5 (c) shows the final binarization result of the contrast image in Figure 5 (b).

3.2 Inpainting Task

The inpainting technique chosen is part of the FoE framework, which was introduced by Roth and Black in (Roth and Black 2005). While in (Roth and Black 2005) priors are used for the modelling of natural images, we utilize image models that are trained on handwritings. Despite this difference, the learning and inpainting tasks are similar to the work by Roth and Black and this section provides a brief overview on the FoE framework.

**FoE Definition**

An MRF is an undirected graphical model, which describes a graph $G = (V, E)$. $G$ consists of nodes $v \in V$, which are connected through edges $e \in E$, whereby the edges model the relationships between the nodes. The nodes in the FoE framework correspond to pixels of an image $x$. A clique $c \in C$ is a subset of neighboring nodes. The probability density of $x$ is defined by:

$$p(x) = \frac{1}{Z} \prod_{c \in C} f_c(x_c)$$

where is a normalization constant, and is the potential function for the clique. MRFs can be categorized into pairwise and high-order MRFs, depending on the clique size. Pairwise MRFs have a clique size of two, whereas high-order MRFs have a larger clique size. The FoE model is such a high-order MRF that models long-range correlations in images. The FoE models, which are utilized in this work, have a clique size of 3 x 3 pixels. Roth and Black have proposed to define the MRF potentials with the Product of Experts (Hinton 2002) representation:

$$\phi(I^T x_k; \alpha) = \left(1 + \frac{1}{2} (I^T x_k)^2\right)^{-\alpha}$$

The learned filter is projected on the image patch. The scalar is also learned from the training database. It is notable that the potential functions used do not depend on the position on the graph, which makes the FoE prior translation invariant.

**Training**

The training procedure is based on the Maximum Likelihood (ML) approach and hence the likelihood is maximized w.r.t. to the
learned parameters. Unfortunately there is no closed form solution for the ML existing and the optimal parameters must be approximated: A gradient ascent is performed on the log-likelihood of the training set. The training set does not consist of entire images, but instead of $15 \times 15$ patches for the purpose of reducing the training time. The FoE models, which are used in this work, have been trained on $20000$ image patches, as it is suggested by Roth and Black. In the original algorithm, those patches are randomly extracted from an image. In our work the patches are also sampled randomly from UV images, but a patch is only added to the training set, if it contains at least a certain number of foreground pixels. Thus, the FoE model is trained on strokes instead of the background.

**Inpainting**

The inpainting formula is based on the Maximum A Posteriori. In general, the posterior depends on the likelihood and on the prior. However, for image inpainting it is assumed that no observation is made inside the inpainting domain and that the observed image outside the inpainting domain is equal to the real data. Hence, the likelihood term, which models the observation process, can be dropped and the restoration depends just on the prior of the input image. In an iterative process the gradient of the prior defined in the following equation is added to the interim result:

$$x^{(t+1)} + \eta M \left[ \sum_{i=1}^{N} \sigma_{i} \psi \left( l_{i} \ast x^{(t)} ; \alpha_{i} \right) \right],$$

where $\sigma_{i}$ is the gradient of $\psi$. The filter is the mirrored version of the filter $\sigma$, is the inpainting mask, is the user-defined inpainting rate and is the number of iterations. The results, which are presented in the next section, are gained by setting $\eta = 100$ and $\alpha = 2300$. We initialize the unknown regions with the dominant background color, since we observed that this measurement prevents an occurrence of smearing artifacts.

4. **Experimental Results**

In the following experiments the Peak Signal-to-Noise- Ratio (PSNR) is used to measure the similarity between a reconstructed image and the corresponding ground truth image. The PSNR is defined as:

$$PSNR = 20 \log_{10} \left( \frac{255}{\text{MSE}} \right),$$

where MSE is the Mean Square Error (MSE). The system is applied on synthetic and real world data in order to illustrate its capabilities.

4.1 **Experiments on Synthetic Handwriting Data**

This experiment is conducted on modern handwritings, which have been randomly extracted from the IAM handwriting database (Marti and Bunke 2002). The FoE prior used has been trained on different words, taken from the same database. The test set consists of 1000 words, which are overlapped by other words with similar sizes. Those overlapping words are also randomly chosen and they define the inpainting masks. The test addresses the issue, which kind of overwritings can be restored in an appropriate manner. Therefore, the overwritings are grouped into 8 groups, based on their average stroke width. The PSNR values, which are achieved inside the inpainting domain, are presented in Figure 6.
The blue bar depicts the PSNR value gained by the FoE inpainting approach and the green bar denotes the PSNR value that is achieved by simply filling the inpainting domain with the dominant background color. It can be seen that the performance of the FoE algorithm decreases with respect to the average width of the mask regions. The inpainting process is often evident, if large gaps - with an average mask width of about 7 pixels or more - are restored, since the inpainting regions are blurred out and the mask centers are not altered adequately. One restoration example, which shows the recovery of the synthetic text data, is given in Figure 7.

5.2 Palimpsest Restoration

The performance on real world data is analyzed by means of manually created ground truth images. Four portions belonging to different palimpsest folios have been extracted from the Archimedes palimpsest. The corresponding ground truth images have been created based on transcriptions, which are provided by the Archimedes Palimpsest Project (Netz, Acerbi and Wilson 2004). It turned out that the resolution of the photographs, which is 700 Dots per Inch, is too large for an adequate filling of the occluded regions, because the mask centers are not altered sufficiently. Hence, it was necessary to downsize the test panels by a factor of four. The PSNR values that are gained by the recovery system are provided in Table 1.

The similarity values inside the inpainting domain are given in the first row of the table. In order to enable an analysis of the entire recovery system, which consists also of the mask generation step, the second row of Table 1 shows the PSNR values of the entire test panels. Although the similarity values inside the inpainting domains are higher than the PSNR values gained on synthetic data, it has to be mentioned that the contrast between text and background is smaller on the real world data. Hence, the MSE of the panels utilized is more bounded compared to the synthetic data. One example for the palimpsest recovery is given in Figure 8. It can be seen that the system is capable of restoring the occluded regions partially. While some characters are successfully recovered, others are not restored in an adequate manner, because the inpainting regions are blurred out. Furthermore, noise is partially enforced during the inpainting process. This noise emerges if the known neighboring regions have a strong background variation. If an overwriting is not entirely identified and the

<table>
<thead>
<tr>
<th>Panel</th>
<th>40 recto</th>
<th>48 recto</th>
<th>58 verso</th>
<th>99 verso</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpainting domain</td>
<td>23.35 dB</td>
<td>24.00 dB</td>
<td>24.31 dB</td>
<td>26.90 dB</td>
</tr>
<tr>
<td>Entire panel</td>
<td>27.29 dB</td>
<td>27.83 dB</td>
<td>28.88 dB</td>
<td>29.94 dB</td>
</tr>
</tbody>
</table>

Table 1. Performance of the Palimpsest recovery.
overwriting vestige is bordering the inpainting domain, it is likely that not existing structures are introduced.

5. Conclusions

This work is concerned with the multispectral imaging and the automated recovery of palimpsests. The proposed system for the palimpsest recovery uses an inpainting approach in order to restore occluded handwriting regions. The inpainting method chosen is part of the FoE framework and allows for an offline learning of handwriting statistics. The technique has the drawback that it is limited to narrow domains and hence it was necessary to downsize the panels utilized. A further disadvantage of the system is the fact that noise is propagated into the inpainting domain and that large areas are blurred out. The models, which are used in this work, have a size of 3x3. We are planning to evaluate the recovery performance of models with a larger clique size, since Roth and Black have shown that 5x5 FoE priors achieve a higher performance than smaller models. Furthermore, we will compare the performance of the FoE approach to the performance of a textural inpainting technique, because such algorithms are not limited to narrow inpainting regions (Wexler, Shechtman and Irani 2007). Hence, a downsampling of the palimpsest photographs could be avoided, which would meet the requirements by scholars, who prefer to analyze high resolution images (Fischer and Kakoulli 2006).

Acknowledgements

This work was supported by the Austrian Science Foundation (FWF) under grant P23133.

References


Wexler, Y., E. Shechtman, and M. Irani. 2007. “Space-
Multispectral Image Analysis of a Censored Postcard from 1942

Florian Kleber, Fabian Hollaus and Robert Sablatnig
Vienna University of Technology, Austria

Abstract:
Multispectral imaging comprises the analysis of objects within different wavelengths to reveal features which cannot be seen by the human eye. In the field of art conservation (e.g. color measurements) and investigation (e.g. analysis of underdrawings in paintings) multispectral imaging techniques, such as Infrared/Ultraviolet Reflectography and Ultraviolet Fluorescence, have been applied. The same methods can be used in historical document analysis to reveal palimpsests and to enhance the legibility of ancient manuscripts. Within this paper the question is raised, if a Jewish postcard from 1942 censored by the Nazi Party, can be made legible by applying multispectral imaging. Hence, this paper gives an introduction to multispectral imaging techniques and its results in historical document analysis.

Keywords:
Multispectral Imaging, Document Analysis, Censored Text

1. Introduction

Censorship of parts of written text was and is a common practice in totalitarian regimes; it is used to destroy information not appropriate to the political power. For historic reasons a recovery of the censored text is of interest to analyze on the one hand the text and on the other hand the censors. Within this paper the question is raised, if a Jewish postcard from 1942 censored by the Nazi Party, can be made legible by applying multi-spectral imaging.

Fig. 1 shows images of the postcard with the blackened text at the reverse side. The author of the postcard, Egon Heysemann, was born on November 25, 1925 in Flatow, a small town in West Prussia. Soon after the Nazi rise to power, the life of Jews in this province became unbearable and many moved to Berlin. In Berlin, Egon Heysemann was placed in the Berliner Auerbach orphanage, which opened its doors to the refugees. There he became a close friend of Rolf Rotschild, the addressee. In 1939 Rolf Rotschild succeeded in escaping Germany with a children transport to Sweden, whereas Egon Heysemann was deported to the East on March 28, 1942. On April 1 he arrived in the ghetto Piaski, Poland. On April 9, he sent a postcard to his friend in Sweden, in which he reported on the deportation and asked for help. The censor made unreadable the details of the trip and the situation in the ghetto that served as a transit camp on the way to the death camps. The date and place of Egon Heysemann’s death are unknown; this postcard is the last known word received from him. Barbara Schieb from the German Resistance Memorial Center Berlin and the current owner of the postcard, Walter Frankenstein, entrusted the postcard to the authors for an attempt at text recovery with the help of scientific methods.

As a preprocessing step a laser cleaning (Pentzien and Kautek 2005; Krüger et al. 2008; Mäder et al. 2007) of the blackened text areas has been performed by the BAM Federal Institute for Materials “Research and Testing”.

Multi-spectral imaging comprises the analysis of objects within different wavelengths to reveal features which cannot be seen by the human eye. The human eye is sensitive to a wavelength range of appr. 400 (blue) to 700 nm (red). Filters (defined by their spectral transmittance) or Light Emitting Diodes (LED) can be used to investigate only a predefined

Corresponding author: kleber@caa.tuwien.ac.at
Multispectral Image Analysis of a Censored Postcard from 1942
Florian Kleber, Fabian Hollaus and Robert Sablatnig

spectral range. Since optical filters degrade the images spatial resolution of an optical system, narrowband light sources such as LEDs are preferred. The use of narrowband light sources additionally reduces the amount of incident light to the painting/manuscript which follows the conservation goal to minimize the incident light dose (Christens-Barry 2007). By using digital sensors, which have an extended responsivity, non-destructive measurements can be performed. InfraRed (IR) rays have the property that they are less scattered and therefore can penetrate materials which are opaque, if VISible (VIS) illumination is used (Mairinger 2003). Nondestructive methods like IR reflectography (Mairinger 2003) are used by multispectral imaging systems to visualize underdrawings in paintings (Hain et al. 2003) and to differentiate between palimpsest texts and the newer text. UltraViolet (UV) fluorescence imaging can be used to enhance the contrast of palimpsests, since “old paint or varnish layers emits more fluorescence light comparing to newly applied materials (repainting or retouching area)” (Hain, Bartl and Jacko 2003). Fluorescence is luminescence that is the emission of one or more photons by an atom (or molecule) that is caused by absorbing a certain amount of electromagnetic radiation. A detailed description of imaging techniques applied to historical documents is described in (Miklas et al. 2008; Easton et al. 2003).

The paper is structured as follows: The following section explains the acquisition setup used for multispectral imaging. Section 3 explains the image enhancement techniques applied to the multispectral images. The restoration results are shown in Section 4. Finally, concluding remarks are given in the last section of this work.

2. Multispectral Acquisition Setup

The acquisition setup used consists of a Near InfraRed (NIR) camera (Hamamatsu C9300-124), and a traditional RGB camera (Nikon D2Xs) mounted next to it. A linear unit allows for a constant shift between the two cameras. The illumination consists of 2 Eureka!Light (Equipoise imaging, Archimedes project (Easton, Knox and Christens-Barry, Multispectral Imaging of the Archimedes Palimpsest 2003)) LED panels mounted on tripods, which allow to use 13 different wavelengths. Additionally 4 panels with white LEDs (color temperature 5600 Kelvin) are mounted left and right of the two Eureka!Light
panels. Additionally a filter wheel, with 8 different filters, is mounted in front of the Hamamatsu camera. The filters used are summarized in Table 1. Currently only 2 filters for UV fluorescence and UV reflectography are used, whereas the others have been replaced by the LED lighting system.

Fig. 2 illustrates the acquisition setup. The entire system is designed to be transportable with a setup time of appr. 1 hour. Table 2 summarizes the nominal (given by the manufacturer) and peak frequencies (most intense wavelength in the emission spectrum) of the used LEDs.

Color images (white LED panels) and UV fluorescence (UV illumination, 365 nm) images are captured with the Nikon camera, which provides a resolution of 4288 x 2848px.

UV fluorescence/reflectography, and all other wavelengths provided by the Eureka!Light are acquired by the NIR (Hamamatsu) camera. A recording area of appr 150 x 110 mm results in a resolution of 4000 x 2672px. As a result the SP 400 and LP 400 filters have to be used.

### Table 1. Description of the multispectral images containing the channel number, the filter type and the methodology of the image acquisition. LP depicts a long pass filter, SP a short pass filter and BP is a band pass filter.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Filter type</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SP 400</td>
<td>UV reflectography</td>
</tr>
<tr>
<td>2</td>
<td>LP 400</td>
<td>Visible light, UV fluorescence</td>
</tr>
<tr>
<td>3</td>
<td>BP 450</td>
<td>Visible light</td>
</tr>
<tr>
<td>4</td>
<td>BP 550</td>
<td>Visible light</td>
</tr>
<tr>
<td>5</td>
<td>BP 650</td>
<td>Visible light</td>
</tr>
<tr>
<td>6</td>
<td>BP 780</td>
<td>Visible light</td>
</tr>
<tr>
<td>7</td>
<td>LP 800</td>
<td>IR reflectography</td>
</tr>
<tr>
<td>8</td>
<td>RGB</td>
<td>Visible light, IR (no filter)</td>
</tr>
</tbody>
</table>

### Table 2. Nominal and Peak Frequencies of the Eureka Light LED.

<table>
<thead>
<tr>
<th>Nominal Frequency [nm]</th>
<th>Peak Frequency [nm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>365</td>
<td>368</td>
</tr>
<tr>
<td>450</td>
<td>444</td>
</tr>
<tr>
<td>465</td>
<td>470</td>
</tr>
<tr>
<td>505</td>
<td>494</td>
</tr>
<tr>
<td>535</td>
<td>518</td>
</tr>
<tr>
<td>570</td>
<td>599</td>
</tr>
<tr>
<td>625</td>
<td>640</td>
</tr>
<tr>
<td>700</td>
<td>706</td>
</tr>
<tr>
<td>780</td>
<td>771</td>
</tr>
<tr>
<td>870</td>
<td>857</td>
</tr>
<tr>
<td>940</td>
<td>927</td>
</tr>
</tbody>
</table>

3. **Image Analysis/Enhancement**

The multispectral scan of the postcard can be interpreted as the linear mixture of several patterns, e.g. the handwritten text, the cardboard, the blackening and so forth. We have applied two spectral unmixing approaches, in order to separate the handwriting from the remaining patterns. These patterns are also named sources and so-called blind source separation (BSS) methods aim at separating the sources from their mixtures. Several authors (Rapantzikos and Balas 2005; Salerno et al. 2007; Easton and Knox 2004; Tonazzini et al. 2004) have proven the usefulness of BSS methods for the recovery of degraded handwritten characters.

The two unmixing techniques, which have been utilized in this work, are named Principal Component Analysis (PCA) and Independent Component Analysis (ICA). PCA is a statistical...
procedure, which transforms the correlated data $x$ into the uncorrelated data $y$. The data is projected on a set of orthogonal axes, whereby the variance of the data is maximized (Rapantzikos and Balas 2005).

In contrast to PCA, the ICA approach (Hyvärinen and Oja 2000) finds a linear transformation that maximizes the statistical independence of the sources. A further contrast to PCA is that the basis vectors of the ICA transformation are not orthogonal. The interested reader is referred to (Hyvärinen et al. 2001) for details on the ICA approach.

4. Results

The results of the MultiSpectral (MS) image acquisition and the BSS techniques applied to the postcard are shown in the following section.

Fig. 3 shows a photograph of the postcard, which is illuminated with 570 nm. A detailed view of the region marked by the green rectangle in Fig. 3 is given in Fig. 4, showing the region imaged at 10 different wavelengths. Like (Stuart 2007), we noticed that the structure of the paper is recognizable in the UV reflectography image.

The first four images in Fig. 4, which were imaged at lower wavelengths, exhibit more background noise than the remaining images, which were photographed at higher wavelengths. We observed that some fractions of characters are located at the censored regions - like the vertical stroke in the upper right corner - are most visible at 570 nm or 625 nm. However, only vestiges of the censored text are visible in these images and the censored parts of the writing remain unreadable. Starting at a wavelength of 700 nm, it can be seen that an enlargement of the wavelengths leads to an increase in the reflectivity of the handwriting and hence to a reduced visibility. The text totally disappears if the postcard is illuminated with 940 nm IR light.
We applied the ICA and the PCA technique to the scan of the postcard to separate the censored text from the blackening. The distribution of the eigenvalues of the covariance matrix is given in Fig. 5. The first eigenvalue includes more than 95%. We decided to use the four largest eigenvalues for the ICA transformation. The PCA-transformed images are given in Fig. 6 and the outputs of the ICA transformation are presented in Fig. 7.

It can be seen that the third output of the PCA approach in Fig. 6 contains mainly the handwriting, while the handwritten text in the first output of the ICA technique has the greatest background contrast, compared with the remaining images in Fig. 7. The ICA output shows fewer portions of the vertical lines than the PCA result does. If we compare the ICA result with the unprocessed images in Fig. 4, we can see that vestiges of characters are recovered by the spectral unmixing technique. However, the restoration result is modest, since a decipherment of the censored text is still impossible.

5. Conclusion

Multi-spectral imaging was applied to the blackened parts of postcard. The multi-spectral system makes use of LEDs, which enabled a non-invasive analysis of the postcard at 12 different wavelengths ranging from UV to NIR. The multi-spectral acquisition did not achieve the desired legibility improvement, since only vestiges of characters are recognizable and these fractions are already best discernible under visible light. We applied two spectral unmixing techniques to the multi-spectral scan in order to separate the writing from the surrounding patterns and hence to enhance the censored text. By using these approaches, the contrast of some character portions was increased, but the censored text remains unreadable.

Acknowledgements

We would like to thank Walter Frankenstein and Barbara Schieb (German Resistance Memorial Center, Berlin) for permission to publish the results of scientific investigations of the postcard. This work was supported by the Austrian Science Foundation (FWF) under grant P23133.

References


Semantic Web Technologies Applied to Numismatic Collections

Ethan Gruber
*American Numismatic Society, USA*

Sebastian Heath
*New York University, USA*

Andrew Meadows
*American Numismatic Society, USA*

Daniel Pett
*Portable Antiquities Scheme, British Museum, UK*

Karsten Tolle
*University of Frankfurt, Germany*

David Wigg-Wolf
*Römisch-Germanische Kommission, Germany*

Abstract

This paper discusses Nomisma.org, a collaborative effort to define numismatic concepts within the linked open data environment. Nomisma.org, created in 2010, defines thousands of concepts, represented by XHTML+RDFa available at stable URIs. The paper also includes test cases for the integration of Nomisma.org URIs into the application architectures at three organisations: why the institutions chose to adopt these concepts and what can be gained by their integration.

Keywords: Linked Data, Numismatics, Museums, Databases

1. Introduction: Why Numismatics?

Numismatics and the projects undertaken by its practitioners provide a near perfect illustration of the potential for a Linked Data approach to transform, simplify and enhance workflows and outputs. The basic activities undertaken by numismatists may be broken down into four main types:

Hoard Studies. Hoards are groups of coins buried together in antiquity, and offer the ability to assess relative chronologies of coins, as well as to analyse patterns of coins circulation.

Individual finds. The finds of individual coins in controlled excavation can often be diagnostic for dating archaeological contexts.

Coin finds, whether in controlled excavation or not, can also contribute to a broad analysis of economic activity across time and space.

Collection catalogues. The major public coin collections in the US and Europe alone contain more than a million specimens of ancient coinage, each of which must be described for collection management and educational purposes. Such coins often bear provenance information, which may link them back to a hoard or individual findspot.

Type Corpora and studies. The basic work of defining coinage through the creation of taxonomies for different series of coinage provides a standard reference system for the cataloguing of coins in all contexts. The establishment of such taxonomies relies,
however, on the assembly and description of vast data sets, from collections, hoards and individual finds. Between these four ‘pillars’ of numismatics there exist a series of functional links that may straightforwardly be described within a Linked Data environment (Fig. 1).

The ability to link and query large bodies of data holds out significant potential for the publishers of catalogues of material to simplify the task they face. But it also offers the possibility to collection managers and publishers to make their material more visible to the taxonomic projects that rely on the accumulation of data and which provide structure to the discipline. Furthermore, the application of Linked Data principles from numismatic material beyond the field to numismatics to other resources in the fields, for example, of geography, prosopography or iconography, allows the incorporation of numismatic material into far broader fields of inquiry.

2. Nomisma.org

Nomisma.org, hosted by the American Numismatic Society since 2010, is an effort to publish disciplinary-specific and stable http-based URIs for numismatic concepts so as to promote interoperability between numismatic collections and projects as well as links to and from other fields of study. While this paper shows that Nomisma.org is an “up and running” project, it does remain very much under development in terms of what it includes and what are the best approaches to ensure widespread adoption within the field of numismatics. That said, its current manifestation is deeply informed by the both the explicit principles and implied best practices of “Linked Data” (Berners-Lee 2006). In particular, Nomisma.org publishes semantically clear URIs, intends that those URIs are permanent, provides access to both human readable and automatically-parsable versions of each RDF resource, and also provides access a single Creative Commons licensed download describing the entire vocabulary. To distill that sentence down to its essence, Nomisma.org promotes permanence for numismatic concepts and the availability of reusable data when the the URIs for those concepts are referenced.

But with that slightly technical introduction in place, it is worth discussing why such a set of stable URIs is useful in the first place. As a discipline, numismatics has a long history during which it has developed a set of terms that have demonstrated their worth for describing coins and other related monetary items. For the purposes of this paper, a coin can be defined as any object - usually metal, small and round - struck by an official entity to a standard weight and appearance in order to facilitate payment in the context of economic transactions.

It is likely that all readers know what coins are and that their experiential definition is sufficient in this context. What may be surprising is that the same set of terms is useful in describing all coins from the 7th century BC onwards as produced by many historical and contemporary states and cultures around the world. In general, coins have an obverse and a reverse - “heads” and “tails” in common parlance. Accordingly, Nomisma.org defines the URIs http://nomisma.org/id/obverse and http://nomisma.org/id/reverse. The writing on a coin is of particular interest and “legend”
is the long standing term-of-art meaning all words, whether abbreviated or not, found on a coin so that http://nomism.org/id/legend allows that information to be clearly marked. More specialised terms include ‘axis’, which is the angular relationship between the obverse and reverse of a coin. Most modern coins have an axis of 6 because the reverse image is upside down in relation to the obverse portrait. The URI http://nomisma.org/id/axis can be used to capture this concept.

The terms introduced above are useful for describing coins, that is individual numismatic items. Nomisma.org also defines more abstract concepts that appear in the field. A hoard is a group of coins concealed together, often with the intention of later retrieval. Ideally, the findspot of a hoard is known, its date of deposit is determined by the date of the latest coin within it, and the origins of all its coins indicate a network of economic interaction. In order to express this set of concepts and relationships using http-based URIs, Nomisma first establishes the URI nm:hoard to indicate that a resource describes such a group of coins. http://nomisma.org/id/igch1546 is the URI for a description of a hoard itself, in this case as described in the standard reference work An Inventory of Greek Coin Hoards (IGCH). (Thompson et al. 1973) Within that hoard, each mint - here meaning the ancient city responsible for the coins production - is marked by a URI pointing to the description of that city. While further information is indicated, establishing the relationship between a hoard and the mints represented in it is sufficient to make a map of the hoard, and also for each mint to make a map of those hoards containing its coins (Figs 2-3). Such mapping relies on knowing the latitude and longitude for each hoard as well as for each mint; information that Nomisma.org supplies for such resource types. Descriptions of the hoards in IGCH were among the first resources added to Nomisma.org and it is likely that the markup for this class of resource will be further developed in response to the work of the Coin Hoards of the Roman Republic project described below.

It is important that the format used to describe each resource is XHTML+RDFa 1.1. “RDFa” - or RDF in attributes - is a format for embedding RDF in xml based documents such as XHTML. (McCarren 2012) “RDF” in turn is the “Resource Description Format”, which has as its underlying data model tripartite statements known as “triples”. A triple is formed by a subject, predicate and object. An example is:

```
http://nomisma.org/id/igch1546

http://www3.org/2003/01/geo/wgs84_pos#latlong"37.97472223.7225"
```

While a full introduction to RDF is beyond the scope of this paper, some focus on the choice of RDFa is relevant. Nomisma.org adopted XHTML+RDFa as its base representation because it allows a single representation to be both human and machine readable. RDFa is a W3C endorsed standard with excellent tooling
support. Accordingly, it is possible to confirm that the information Nomisma.org intends to record for each resource it defines is actually parsable by third parties. The same applies to the choice of XHTML. In the context of Nomisma.org, XHTML also stands as a W3C-supported standard with excellent support. Of course, the primary tool for accessing XHTML is the browser, which is essentially ubiquitous in today’s computing environment. Accordingly, it is likely that the combination of XHTML and RDFa will mean that Nomisma.org’s definitions of numismatics concepts are readily usable both now and far into the future.

As noted in the first paragraph of this section, Nomisma.org is in early stages of establishing stable URIs for numismatic resources. Nonetheless, both within the context of Nomisma.org itself and by third-party initiative, patterns of usage for Nomisma.org URIs are emerging. Within Nomisma.org, the description of coins types, such as those based on the reference works Roman Republican Coinage (Crawford 1975) and Roman Imperial Coinage best demonstrate how URIs can be used in constructing descriptions of well-known coins. http://nomisma.org/id/rrc-525.4a is a coin type within the RRC series. The RDF embedded in this resource is shown in Figure 4. The ontology represented is particular to Nomisma.org, but coin type records will alternatively be provided in a model that conforms to The British Museum’s CIDOC-CRM template for coins. Note that where possible, the RDF refers to other Nomisma.org URIs. Furthermore, if one were to dereference the URI for the mint of this coin type, in this case nm:rome, one would find a further reference to the URI to the Wikipedia page for “Ancient_Rome” and to the identifier of this site within the Pleiades Project. Pleiades is an NEH-funded project to assign identifier to places relevant to the study of ancient geography. By providing links from coin types to Pleiades IDs, Nomisma.org enables interoperability with projects that lie outside the scope of numismatics. The following sections of this paper demonstrate current third-party uses of Nomisma.org URIs.

3. Early Steps Toward Nomisma Integration

Nomisma.org is now two years old, but it has been gaining traction within the numismatic community over the previous six to eight months, beginning with a meeting held at the British Museum October 2011. In this meeting, the European Coin Find Network and Portable Antiquities Scheme committed to collaborating with the American Numismatic Society to associate their records with Nomisma URIs as an early step to exposing their data to the wider linked data universe. This section of the paper details the processes undertaken by the three organisations to incorporate Nomisma into their application architecture.

3.1 Current Work at the European Coin Find Network

In a joint project, Databases and Information Systems (DBIS) and the Römisch-Germanische Kommission (RGK) are investigating the logical integration of different European coin find databases using ontologies. The two main benefits of ontologies in our view are: a) problems are lifted to a content level and are not additionally hindered by technical issues, and b) ontologies provide the possibility of viewing data from a new perspective – the links between the artefacts can be visualised and are not hidden in flat tables.
In an earlier pilot project, InterFACE (Internet Portal: Finds of Ancient Coins in Europe), involving DBIS and the Mainz Academy, a meta-portal was set up containing an integration layer based on a global ontology to access three online databases in Frankfurt (NUMIDAT-WEB), Utrecht (NUMIS) and Vienna (dFMRÖ) (Lehmann and Varughese 2008). Queries were defined via a Web frontend. These queries were based on a global ontology and then translated to corresponding SQL statements for the different databases.

Mapping between the databases and the global ontology of the meta-portal’s integration layer was not done directly, but by means of a kind of mediator ontology—similar to the data model of the individual databases—which was first defined for each of the sources. This turned out to be a particularly efficient method.

In a next step the lessons learned are to be applied to a meta-portal for the European Coin Find Network (ECFN), again accessing databases of coin finds, but also incorporating aspects of Nomisma.org. There are several levels and ways in which Nomisma can be employed: in the individual source databases, in the integration layer of the meta-portal, or in the results.

Within the InterFACE project the relational databases being accessed had been designed with extremely different modelling principles – in one case all information was stored in one large table – while other databases had a high degree of normalisation. These modelling principles also have an impact on the data quality of the individual databases. One large table which does not have the benefit of referential integrity can result in various entries for the same thing, e.g. by a typing or import error. The result might be conflicting entries for material such as: a) “silver”, b) “ar” – Latin argentum, abbreviated, or c) “sliver” – the result of a typing error. Introducing Nomisma-IDs (the ID for silver is: http://nomisma.org/id/ar) to the source databases automatically would solve this problem by providing external referential integrity.

However, our current main concern is that of mapping between different relational database sources. Besides the quality problem of the individual databases, there are also language problems, e.g. “silver” (English) vs “Silber” (German). Nomisma-IDs can act here as a lingua franca, since both point to the same ID. This can simplify the mapping and should facilitate the use of partially automatic mapping processes.

Finally, the results from the meta-portal can be enriched by using Nomisma-IDs to include information drawn from Nomisma or linked resources that was previously not included within the databases connected to the meta-portal. To stick to the example of silver; it would be possible to use the Nomisma-ID in order to access the information that “Argentum” is an alternative label for silver, and “AR” is its abbreviation in Nomisma.org. Beyond this we could learn that it is classified as a material.

An additional long term benefit would be that reasoners could scan all of this ontological information. Based on defined rules, they could then be used to search for potential errors revealed by objections, thus increasing data quality, or to find new deductions that had not been apparent before.

However, for the time being the focus will be on setting up a meta-portal for ECFN. Within this portal Nomisma-IDs will be used for the internal ontology, and we will experiment with them to see how to simplify and optimise the process of mapping different databases, and how to support this with various tools. For example, DBIS developed a tool to this purpose, Integration with Ontologies (IwOnto) (Tolle and Bachmann 2012). IwOnto allows mapping a relational database to an existing ontology and to export database instances within the ontology. In this way IwOnto can be used to exchange data between databases.
to perform their physical integration. The resulting ontology containing the instances can also be used by Semantic Web tools such as reasoners, or to visualise data in a more sophisticated manner, e.g. using the Protegé plugin Jambalaya (http://www.thechiselgroup.org/jambalaya).

3.2 The American Numismatic Society: Two Test Cases

Nomisma.org is the cornerstone for all current and future projects at the American Numismatic Society. The role of Nomisma in the ANS project, Online Coinage of the Roman Empire (OCRE), was presented at CAA 2012 with the paper “Linking Roman Coins: Current Work at the American Numismatic Society.” OCRE was neither the first, nor will be the last ANS project to assimilate Nomisma concepts into its core architecture. This section of the paper will illustrate two test cases of the incorporation of Nomisma URIs to facilitate interlinking. One project is Mantis, the ANS’s newest collection interface (http://numismatics.org/search/; the other is a catalogue of Roman Republican Coin Hoards. Both of these projects are built on top of the Numishare framework, an open-source numismatic collection management system, which is discussed in greater detail in the aforementioned paper.

Mantis

The American Numismatic Society first endeavored to integrate Nomisma URIs into Mantis in the beta-testing phase of the project prior to its official release in April, 2011. As of June 2012, only URIs for mints are attached to coins, which are primarily of Greek, Roman, and Byzantine origin. Linking coins to mints enables the ingestion of geographic coordinates provided by the Nomisma RDF into Mantis’ search index for mapping applications. Objects in Mantis will eventually link to Nomisma for other numismatic concepts, but this process of integration is gradual and related directly to the curation and standardisation of data within the ANS’s Filemaker database.

The digitisation of the ANS collection began in the 1980s with a DOS-based database, which was eventually migrated to Filemaker Pro within the last decade. Over this period, many curators and interns have made contributions to the database, although there was minimal adherence to controlled vocabulary or enforcement of quality control until very recently. Furthermore, encoding practices varied from department to department, which is to say, identical database fields may have held different semantic meanings to different departments within the Society. This can be best illustrated with the database’s three geographic fields: mint, region, and locality. An American coin may use mint: Philadelphia, locality: Pennsylvania, region: United States. A Greek coin from Olynthus (mint), Macedon (region), Chalcidian League (locality). While the mint field typically captures the coin’s city of origin across all departments, the use of region and locality differ. In ancient coins, regions are purely geographic, but the field is used to capture the name of the nation-state in more modern specimens. Locality is intended to be a hierarchical child of region. The Chalcidian League is an authority responsible for the production and distribution of coins, not a geographical identifier. Hence, the shortcomings of the current database framework are readily apparent.

To facilitate the linkages between the ANS collection and Nomisma.org upon publication of the Filemaker data into Mantis, a “lookup table” was created in a Google Spreadsheet. This spreadsheet contains columns for mint, region, and locality from Filemaker and a reference to the Nomisma URI for each particular combination of fields. This enables coins encoded with differing mint labels of “Antioch” and “Antiocheia” to be linked to http://nomisma.org/id/antiocheia_syria. Similar lookup tables can be created to regularise other types of numismatic concepts.
in Mantis and will provide the capacity to migrate from Filemaker into a framework that supports more sophisticated controlled vocabulary integration at some point in the future. While Mantis supports querying by Nomisma and Pleiades URIs for mints, the ultimate goal is to enable the ANS collection to be queried among other collections, regardless of language, by linking objects even more closely to Nomisma, as well as other authority services such as Geonames.org and the Virtual International Authority File (VIAF.org). The Filemaker regularisation may require several years’ effort, but the benefit is enormous to the broader numismatic community.

**Roman Republican Coin Hoards**

As discussed earlier in this paper, Nomisma.org provides stable URIs for coin types from two Roman numismatic catalogues: Roman Republican Coinage (RRC) and and Roman Imperial Coinage (RIC). If we consider this coin type metadata delivered through Nomisma to be canonical representations of the information conveyed electronically about the printed reference catalogue, then we may develop tools which leverage this metadata to create an interface that allows users to interact with it.

The American Numismatic Society is currently developing a prototype for delivering coin hoard data online, focusing primarily on those of the Roman Republican era, although these methodologies can be applied to hoards of other periods. Kris Lockyear, an archaeologist at University College London, provided a Microsoft Access database of Republican hoards, which was migrated into an XML-based schema designed to accommodate the many permutations for a hoard record. Variations in descriptive practice range from the numeric totals of very generalised typologies (e.g., authority or denomination) to totals by coin type, or even references to stable URIs of physical specimens available on the Internet (which may include coins in the ANS’s own collection). In this particular project, a coin hoard record includes a list or Nomisma URIs for Crawford’s RRC identifiers — like http://nomisma.org/id/rrc-100.1a — and totals associated with a particular findspot (a modern location identified by a Geonames URI). Numishare, the publication framework for serialising these XML records into HTML (as well as KML for mapping), extracts the metadata for each coin type from the RDF to create a comprehensive display of the record as well as for quantitative analysis (Fig. 5).

The quantitative analysis feature of the project is fairly rudimentary at this stage of the prototype, but will grow in sophistication over time. Interpreting a combination of numeric counts per coin type and the number of distinct typological attributes (e.g., issuers, mints, or denominations) contained within the coin type metadata record, Numishare feeds an array of values into the Highcharts Javascript library to output bar graphs of selected queries. This straightforward method of visualisation is very valuable to researchers. Such charts which once required significant expenditure of time in the pre-digital age can be now be generated nearly-instantaneously.

Like Mantis, the Roman Republican Coin Hoard online catalogue can be queried...
by Nomisma URI, enabling cross-collection searching. Among the most important aspects of the project is the mechanism for referencing URIs and dynamic processing of data hosted and maintained by Nomisma.org, visualizing queries in the form of charts and maps (Fig. 6). The editors of the coin hoard metadata are not burdened with the long-term maintenance of the typological data in Roman Republican Coinage; hoard XML records contain scarcely more than findspot information and a list and count of coin types associated with the hoard, and therefore duplication of data and digital curation labour is minimised. This project is a successful proof-of-concept for other, larger coin hoard recording efforts.

3.3 The Portable Antiquities Scheme

The British Museum’s flagship Public Archaeology project, the Portable Antiquities Scheme (PAS), faces different challenges to many of the organisations that have committed to furthering the uptake of Nomisma. The PAS database (http://finds.org.uk/database) and recording structure has been developed for the recording of all types of singular archaeological discoveries or ‘finds’ dating to over 300 years ago, encompassing objects that range from buckles and brooches of all periods to the most prolific discovery - numismatic material. The PAS also records coin hoards under the remit of the Treasure Act (1996), and with the numismatic material mentioned above this provides a foundation upon which standardisation can be applied.

Numismatics provides over 37% of the total objects (approximately 796,000) recorded by PAS’s 60 members of staff and an army of volunteers (see table 1 for a breakdown). These data include figures for some coin hoards (such as the magnificent Frome Hoard (Booth et al 2011 & Bland et al 2010)), but no means all. A project seeking funding at the moment from the AHRC hopes to fully catalogue and record all Roman coin hoards found in the UK and this will have a significant impact on the quantity of coins available.

Since 2006, the PAS database has implemented strictly controlled vocabularies for the recording and dissemination of information about coins; for example Reece periods, issuer and mint lists, reverse types and Republican moneyers. As detailed in the section relating to the ANS, the PAS also had a period of de-normalised data entry, which has led to expensive (time) exercises to cleanse and enhance old data. It is readily apparent, that many in the numismatic community share the same technical problems, and these have readily been reinvented as more digital projects come online. Ergo, aside from the commonality of problems, this standardised numismatic material provides partners within the Nomisma consortium a solid foundation.

Since 2003 (Pett 2010a, 1), these data have been surfaced for reuse, via a variety of standardised interoperable methods (OAI-PMH - see Blackburn 2005 for early examples
of numismatic OAI use, Z39.50, MIDAS XML and subsequently via json, kml, atom/rss and embedded rdfa) and it is now that the potential of ‘linked data’ is being explored and developed further. These data feed into several large scale Cultural Heritage projects, for example Europeana and Culture Grid, and also to the regional Historic Environment Record.

At the time of writing, the database is being redeveloped to allow for successful integration of Nomisma concepts and URIs as discussed above in Section 2, for example numismatic material provides access to information relating to people (minted for, minted by, depicting), place (place of issue, place depicted, place of discovery) and time (date issued, date found, date commemorating). All of these entities provide a rich tableau from which they can (or will be) linked to third party resources such as dbpedia, geonames or Pleiades## producing enriched (machine readable) records. Detailed maps (Fig. 7) with a variety of layers, showing place of discovery and place of manufacture can be produced and ancillary data (such as podcasts and video) can be introduced to point the researchers to view extra resources. The possibilities for enrichment are manifold and will increase as more people subscribe to this development of a community of practice.

Harnessing these third party tools has led to the development of numismatic resources with stable, permanent, cool URIs (see Sauermann & Cyganiak 2008 for further discussion) which have provided an accessible entry point for learning about numismatic material. Figure 7 depicts the PAS entry for Nero#, displaying data resources drawn from amongst others, Nomisma, dbPedia, VIAF and the PAS database itself (for example, the latest coin records.) An aspiration for these resources, would be the possibility for linking to external numismatic collections that held coins relating to Nero - for example, The British Museum now has an RDF triple store# based around an extended CIDOC-CRM schema, which if linked to Nomisma or dbPedia ontologies, would allow for the retrieval of their’s and other’s data. Thus returning to the introductory text detailing the ability to query large datasets at will and link them through a common paradigm. As detailed above, the ANS as founding member, is committed to implementing the Nomisma concept, and so one could also retrieve their data based around the a simple ‘personal’ resource - http://nomisma.org/id/nero or a combination with a ‘geographical’ resource such as http://pleiades.stoa.org/places/423025 (Rome) or geonames or Yahoo! WOEID#. Harnessing these linked resources would enrich the research potential of a digital resource immeasurably and provide cross-walking between different institutional repositories and aid resource discovery. The PAS model for developing its numismatic resources has been heavily influenced by the collaborators on this paper; aspects such as statistical analysis of coins, emulating the Mantis model, will be introduced, Solr search has been constructed and emulation of the RDFa travails on Nomisma is underway. The URL http://finds.org.uk/romancoins/emperors/emperor/id/12 illustrates a coin page for the Emperor Nero, with reference to and integration from resources such as dbPedia, VIAF, Nomisma, Wikipedia and the internal content management system.

Multilingual numismatics

A challenge that has hindered the PAS’s attempts to share its numismatic database, has been language based, and it is this issue that
the advent of Nomisma might solve. Services such as dbpedia, already provide multi-lingual SPARQL endpoints for querying concepts and entities and this methodology is one that Nomisma can aspire towards. For example, pulling in multi-lingual content from dbPedia can be done via a query as shown below:

```sparql
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX dbpedia2: <http://dbpedia.org/property/>
PREFIX dbpedia: <http://dbpedia.org/>
PREFIX dbpedia-owl: <http://dbpedia.org/ontology/>
SELECT *
WHERE {
  OPTIONAL {?emperor dbpedia-owl:abstract ?abstract} .
  OPTIONAL {?emperor foaf:depiction ?depiction} .
  OPTIONAL {?emperor dbpedia-owl:thumbnail ?thumb} .
  OPTIONAL {?emperor dbpedia2:imgw ?imgw} .
}
```

This SPARQL query retrieves all results for Nero with data for an abstract, a picture, a thumbnail and the width of the thumbnail and the Wikipedia page, however, by appending this statement before the closing curly brace - FILTER langMatches( lang(?abstract), "en") - one can limit searches to just English content. Via these queries, information can be retrieved and linked to for a wide array of information; place of birth, successor, depictions, abstracts etc. Once this multilingual querying facility is implemented, one could create a Europe wide database where all contributors could and should be talking about the same numismatic concept, with no ambiguity or arguing.

### 4. Conclusions

Two meetings held in Frankfurt and Berlin in May 2012 attracted representatives from numerous European numismatic organisations. The databases associated with these organisations are in English, German, French, Slovenian, Dutch, and Romanian. The general consensus among participants of these meetings is that Nomisma.org is a valuable resource for facilitating internationalisation and cross-collection querying. Each sees the shortcomings of their own data systems and is interested in collaborating to grow and improve Nomisma. At the present time, Nomisma RDF contains only English labels, but emphasis will be placed in the coming months and years to introduce labels in other languages, which will enable the goal of internationalisation to become a reality. Nomisma is a cornerstone in the advancement of the discipline of numismatics into the 21st century, and will be paramount in the evolution of the discipline (with respect to both its study and publication) over the next ten to twenty years.

### References


Automatic Coin Classification and Identification

Reinhold Huber-Mörk
AIT Austrian Institute of Technology GmbH, Austria

Abstract:
We investigate coin recognition and classification in a setting with a large number of classes as well as recognition and identification of individual coins of high similarity. Real-world coin image data sets were obtained for the classification and identification tasks. The considered classification task is the discrimination of modern coins into several hundreds of different classes. Identification is investigated for hand-made ancient coins. Intra-class variance due to wear and abrasion vs. small inter-class variance makes the classification of modern coins challenging. For ancient coins the intra-class variance makes the identification task possible, as the appearance of individual hand-struck coins is unique. Modern coins were acquired by a high-speed machine vision system for coin sorting. For ancient coins the setting is more general, images acquired by scanner and camera devices are considered. Methods for automatic image preprocessing and feature extraction are discussed. Results for automatic classification, identification and information fusion are presented.

Keywords:
Image Processing, Coin Classification, Coin Identification

1. Introduction

For modern coins, i.e. machine struck coins, judging systems using electromechanical devices are wide-spread. Those systems are commonly based on measuring weight, diameter, thickness, permeability and conductivity (Davidsson 1996), oscillating electromagnetic field characteristics (Neubarth et al. 1998), and photo- and piezoelectric properties (Shah et al. 1986). Typically, such systems are only capable to discriminate a small number of different coin denominations and are mostly limited to a specific currency. Furthermore, such systems could only be used for machine-struck modern coins.

In general, approaches towards classification of modern coins using image processing are described in various papers and patents. A neural network approach capable of discriminating between 500 Won and 500 Yen coins was published by Fukumi et al. (1992). The so called Dagobert coin recognition system was developed for high volumes of coins and a large number of currencies (Fürst et al. 2003; Nölle et al. 2003). Image binarization followed by area measurement and comparison of coin center and center of gravity was suggested in a patent (Onodera and Sugata 2002). Another system based on the analysis of one side of a coin by transformation of its image into polar coordinates and matching of profiles taken along angle direction was described by Tsuji and Takahashi (1997). A special acquisition device for coins employing colored illumination from various angles was suggested by Hoßfeld et al. (2006). Methods based on matching gradient directions (Reisert et al. 2006; 2007) and color, shape and wavelet features (Vassilas and Skourlas, 2006) were suggested. An approach based on multiple Eigenspaces aims at classification for a large number of classes (Huber et al. 2005).

For ancient coins, i.e. hand struck coins, some publications discussing approaches for classification appeared. Early approaches, which achieved a moderate classification performance, were based on matching of contour and texture features (Van Der Maaten and Postma 2006) or make use of interest
point extraction and matching of local features (Zaharieva et al. 2007). More recently, an approach based on interest points and improved feature description and matching was reported (Arandjelovic 2010). The inherent properties of hand struck coins result in individual features for each coin and a large intra-class variance. Therefore, object classification becomes challenging. However, in contrast to object classification, object identification relies on those unique features which distinguish a given object from all other members of the same class. Results on identification of ancient coins were reported by Huber-Mörk et al (2008), where the combination of shape and local descriptors to capture the unique characteristics of the coin shape and die information was suggested. For ancient coin recognition features from the Scale-invariant feature transform (SIFT) (Lowe 2004) was used and compared to algorithms based on shape matching i.e. a shape context description and a robust correlation algorithm (Zaharieva et al. 2007). From a numismatic point of view, the shape of an ancient coin is a very specific feature. Thus, the shape described by the edge of a coin serves as a first clue in the process of coin identification and discrimination. A shape based method tuned to the properties of ancient coins was combined with matching of local features through Bayesian fusion (Huber-Mörk et al. 2011).

We will review an approach for classification of coins based on an Eigenspace representation (Huber et al. 2005). This approach is well suited for the classification of modern coins and makes additional use of geometric measurements and information extracted from obverse and reverse sides of the coins. Incorporation of geometrical measurements and fusion of coin sides is realized by preselection and Bayesian fusion. Rejection of unknown coins and a discussion on false classification and false acceptance rates vs. false rejection is included. For several reasons, this approach is not suited for ancient coins. For identification of ancient coins we will consider an approach based on local features derived from the coin surface combined with shape features describing the edge of a coin (Huber-Mörk et al. 2008; Zaharieva et al. 2007). Comparison of two coins is done by matching the features derived from contour representations and refined by local features matching. Bayesian fusion of coin sides is studied. In order to discuss the identification performance a discussion of precision vs. recall is included.

The data set for classification of modern coins consisted of approximately 12,000 coins with images of reverse and obverse sides. The data set contained 932 different coin classes. A derived data set was made publicly available as a benchmark by the EU MUSCLE network of excellence (Nölle and Hanbury 2006; Nölle et al. 2006). The data set for identification of ancient coins was provided by the Fitzwilliam Museum, Cambridge, UK and was made publicly available by the EU COINS project (Kampel et al. 2009). The data set consists of 240 coins of the same class with 1200 images of obverse sides and 1200 images of reverse sides acquired by different acquisition devices.

This paper is organized as follows. Section 2 describes coin detection and invariant preprocessing and Section 3 discusses matching based on various feature descriptors. Classification, identification and information fusion is described in Section 4. Results are presented in Section 5 and conclusions are drawn in Section 6.

2. Coin Image Processing

The appearance of coins in 2D images is highly influenced by the lighting conditions and the orientation of the imaged surface. Coins are characterized by a 3D surface and the reflected light into the camera direction is typically a mixture of strong specular and diffuse reflections depending on the placement of camera and light sources, the type of light sources, the coin surface structure, dirt and
abrasion. In order to diminish the influence induced by the lighting conditions a controlled acquisition setup is recommended. Controlled acquisition strongly improves recognition of objects of low intra-class surface variation, e.g. modern coins. Ancient coins are characterized by high surface variation even within a single class, therefore different type and direction of light sources make small patterns on the coin look very different which limits, for instance, the use of local image features for coin recognition. Best practice for acquisition of ancient coins was summarized by Kampel and Zambanini (2008) and Hoßfeld et al. (2006) described a sophisticated system for modern coin acquisition.

2.1 Coin Detection

The separation of an object of interest in an image from background is commonly termed segmentation. Under controlled acquisition automatic intensity thresholding approaches (Sezgin and Sankur 2004) are feasible for modern coins (Nölle et al. 2003). Due to textured background, presence of other objects in the image, inhomogeneous or poor illumination and low contrast, straightforward methods based on global image intensity thresholding tend to fail. In situations, where explicit knowledge on the properties of objects is available, this knowledge can be used to steer segmentation parameters. In our work, ancient coins were localized by thresholding the local intensity range, i.e. the difference between maximum and minimum grey level, in a local window and evaluation of the compactness measure (Zambanini and Kampel 2008). Typically, the shape of modern coins is circular, whereas ancient coins deviate from this shape, but still stay close to a circular outline. Therefore, approaches based on edge detection and application of the Hough transform (Duda and Hart 1972) were applied to modern coins (Reisert et al. 2006) as well as to ancient coins (Arandjelovic 2010), where a modified version of the Hough transform was used.

For a modern coin, such as shown in Fig.1(a), we suggest an edge based technique to segment the coin from the background. The detection of the coin employs a common segmentation approach and works reliably for controlled lighting conditions and relatively clean background, e.g. a moderately dirty conveyor belt. Problems might be caused by very dark coins, i.e. coins which reflect only a small amount of light towards the camera. A multi-stage segmentation procedure is suggested. The outline of the suggested segmentation method

\[ \text{Figure 1} \text{. A modern coin: (a) coin image, (b) smoothed coin image, (c) edge image, (d) labeled image, (e) binary image with convex hull, (f) overlay of segmented image and coin.} \]
is as follows:

- Smoothing of the image to suppress noise and background texture, see Fig.1(b).
- Edge filtering using a Laplacian of Gaussian approach followed by zero-crossing detection (Marr and Hildreth 1980), see Fig.1(c).
- Labeling of the detected regions, see Fig.1(d), and selection of the region with largest bounding box as coin region candidate.
- Form a blob by computing the convex hull of the coin region candidate, see Fig.1(e).

An example of an overlay of the extracted blob onto the input image is shown in Fig.1(f). Coin position and diameter are estimated from the detected blob, which directly delivers access to a translation invariant description.

For ancient coins we employ a measure of compactness $c_t$ related to a threshold $t$ defined as:

$$c_t = 4\pi A_t / P_t^2,$$

where $A_t$ is the area of the region covered by the coin and $P_t$ is the perimeter of the coin. The measures $A_t$ and $P_t$ are obtained by connected components analysis (Sonka et al. 1998) applied to the binary image which is derived from thresholding the intensity range image.

Figure 2 (a) shows an intensity image of an ancient coin, Fig. 2 (b) is the corresponding intensity range image and Figs. 2 (c)-(e) show thresholded images for different selections of $t$ along with calculated values for compactness $c_t$. The image thresholded at the optimal level with highest compactness is given in Fig. 2 (f). Oversegmentation of the coin into several small regions results in small compactness, e.g. compare to Fig. 2 (e).

2.2 Invariance with Respect to Intensity

Apart from illumination dependency the appearance of a coin varies considerably
with respect to its grey values depending on dirt and abrasion. These variations frequently are inhomogeneous. This suggests, even if illumination influence could be neglected, that for recognition purposes grey values by themselves will not give appropriate results. On the other hand, edge information remains more or less stable or at least degrades gracefully. Therefore, we based the feature extraction for coin recognition on edges. In principle any edge detector may be used for this purpose. From our experience the approaches suggested by Canny (1986), by Rothwell et al. (1995) and the Laplacian of Gaussian method (Marr and Hildreth 1980) work satisfactorily. The Rothwell edge detector was used in our experiments, where the parameters were chosen as follows. Sigma was set to 1.5, scaling parameter of the threshold surface was 0.85 and the percentage of non-pixels was 0.6.

2.3 Rotational Invariance

We obtain rotational invariance by estimation of the rotational angle followed by a rotation into a reference pose. Angle estimation is performed for images transformed into polar coordinates. In the polar image shift invariance, corresponding to rotational invariance when mapped back to Cartesian coordinates, is obtained. Rotational invariance for a coin edge image involves cross-correlation with reference edge images. The edge image is mapped from Cartesian to polar coordinates, see Fig.3. The result of cross-correlation between the coin image to be classified and a set of reference images is used to derive class hypotheses. In detail, for both sides of a coin under investigation rotational invariant processing and hypothesis generation proceeds as follows:

- Estimation of coin diameter and center from coin detection and transformation of the coin images from Cartesian to polar coordinates.
- Selection of a set of reference images depending on thickness and diameter measure (if available). Each reference image is associated with a coin class.
- Subdivision of images into n bands along the radius direction in polar space.
- Band-wise cross-correlation of the coin side edge image under investigation with all reference coin edge images in the selected reference set, resulting in cross-correlation values and an estimation for the associated rotation angle for each reference class.
- Ranking of the reference set by the maximum correlation value and generation of a set of hypotheses for the highest-ranking classes.
In our implementation n=16 bands were used in the polar space. Images were scale to 128 x 128 pixels in Cartesian space and became 128 pixels in radius and 256 pixels in angle direction after transformation. The estimated rotation angle is obtained as the average of the estimated rotation angles for all significant bands of a coin. Bands with less than 7 edge pixels or more than 77 were not considered in the angle estimation, as those bands are considered to bear little or not very well structured information.

3. Coin Classification and Identification

Matching for classification or identification is based on edge based features extracted as described in the previous section. In this section, we will discuss Eigenspace matching and shape matching.

3.1 Eigenspace Description and Matching

The Eigenspace decomposition for image analysis was introduced by Sirovich and Kirby (1987) and found numerous applications over the last decades, most prominently in the field of face recognition (Turk and Pentland 1991). We start with the description of the Eigenspace construction employing principal components analysis (PCA).

In the Eigenspace approach, we consider a set of M images $B_i$ to $B_M$. Each image $B_i$ is of size $N \times N$ pixels. The images are reformed into vectors $G_i$ to $G_M$, e.g. by scanning the image line by line. If all pixels of an image are used to produce a vector, each vector $G_i$ has length $L = N^2$. An average vector $Y$ and difference vectors $Y_i - Y$ are calculated and a covariance matrix $C$ is constructed for the set of difference vectors. The Eigendecomposition of the covariance matrix $C$ delivers Eigenvectors and associated Eigenvalues. The Eigenvectors are sorted in non-increasing order depending on the corresponding Eigenvalue. A small number $M'$ of significant Eigenvectors is retained from the ranked Eigenvalues, a common practice which leads to the most expressive features (Turk and Pentland 1991). A weighting factor $\omega_k$ corresponding to the $k$-th Eigenimage is obtained by projection of the image $G$ onto the $k$-th Eigenspace component $u_k$ using:

$$\omega_k = u_k^T (G - Y), \quad k = 1, \ldots, M'.$$

The weights $\omega_k$ are summarized in the vector $\Omega = (\omega_1, \ldots, \omega_M)$ and Eigenspace matching becomes comparison of weights. For the coin recognition task, not the full images are reformed into a vector $\Gamma$, only the interior pixels of the coin are rearranged into the vector $\Gamma$, see Fig. 4.

We did not use grey value images for input to the eigenspace, we used the so-called Eigenhills approach instead (Yilmaz and Gökmen 2000). Eigenhills are derived from application of the PCA to smoothed edge images. We used a 2D Gaussian filter kernel with a sigma of 1.5 to smooth the edge images which are of size 128x128 pixels. Figure 5(a) shows an example of the first 32 Eigenhill images ranked by corresponding eigenvalues and Fig. 5(b) shows the cumulative sum of Eigenvalues depending on the number of Eigenvectors used. Compared to the intensity Eigenspace, where 78% of the
variance is captured by 32 eigenvectors ranked with respect to the corresponding Eigenvalues, the Eigenhill approach achieves 76%. Thus, the Eigenhill approach has similar performance with respect to Eigenvalue preservation while obtaining much better invariance with respect to illumination.

The chosen image size of 128 x 128 pixels was derived from the acquisition resolution, which was 0.05 mm, and typical coin diameters, i.e. the smallest coins are only moderately upsampled. Larger images need to be subsampled to fit to 128 x 128 pixels, which is a tradeoff between accuracy and computational effort of the PCA. In practice, multiple Eigenspaces are constructed. Coins with similar diameter and thickness, if this information is available, are grouped into the same Eigenspace. In our case 494 Eigenspaces were constructed, each of which holds coins of similar dimensions. Depending on the estimated coin dimensions the appropriate Eigenspace is selected and the subsequent decision of the class membership is taken among a smaller set of classes.

3.2 Shape Description and Matching

The shape descriptions of two coins are compared by a linear combination of global and local shape matching. The local matching is derived from the difference of Fourier shape descriptors, whereas matching error between the curves serves as global measure of shape dissimilarity. Our approach of shape comparison is based on a description of the difference between the shape of a coin and the shape of a circle. Therefore, the suggested approach is called deviation from circular shape matching (DCSM).

3.2.1 Shape Description

In order to represent the coin shape, a border tracing on the binary image resulting from segmentation is performed. A list of border pixels is obtained and resampled to n samples using equidistantly spaced intervals with respect to the arc length. Figures 6 (a)-(d) show this operation.

A one-dimensional descriptor, i.e. a curve describing the border, is obtained from fitting the coin edge to a circle and unrolling the polar distances between sample points and fitted circle into a vector. The center $s_c = (x_c, y_c)$ of the fitted circle is derived from the center of gravity and the radius $r$ is the mean distance between the center and all sample points $s_i = (x_i, y_i)$. The 1D representation $S = (d(1),..., d(n))$ is given by:

$$d(i) = \frac{||s_i - s_c||_2 - r}{r}, \quad i = 1, \ldots, n,$$

where $||\cdot||_2$ denotes the $L_2$-norm. The division by $r$ makes the representation invariant with respect to scale. Figure 6 (e) shows the obtained 1D representation.

3.2.2 Shape Matching

In order to compare two shape
descriptions $S_A$ and $S_B$ we denote its Fourier transform coefficients by $F_A$ and $F_B$. The squared distance between the magnitude values of the Fourier coefficients is used as local measure of dissimilarity, i.e.

$$D_L = \sum_{i=v,...,n-u} ||F_A(i) - F_B(i)||^2 / (n - u - v + 1).$$

The lower $v \geq 1$ and upper offsets $u \geq 0$ for the Fourier descriptors are small constants and used to limit errors stemming from imprecise circle fitting and quantization noise.

The global shape matching is obtained from the minimum mean squared error (MSE) for all possible shifts of the 1D shape $D_G = \min_{i=1,...,l} \text{mse}(i)$ using:

$$\text{mse}(u) = 1/n \sum_{i=v,...,n-u} (d_A(i) - d_B(i+u))^2,$$

where $i + u$ might exceed $l$ and modulo addition is applied. The position of the minimum of $D_G$ is related to the rotation angle between the compared coins. The overall measure of shape dissimilarity becomes:

$$D_{AB} = \alpha D_L + (1 - \alpha)D_G,$$

where the weighting factor $\alpha \in [0, 1]$ controls the influence of local and global dissimilarity terms.

In our experiments, we used $n = 256$ contour points, for upper and lower offsets $u = v = 3$ was found appropriate. Variation of $\alpha$ was performed and a value of 0.975 provided the best matching results.

4. Classification and Recognition

A framework for classification and identification based on preselection, classification and Bayesian fusion is presented. For modern coins classification based on Eigenspace representation and prior information and fusion of obverse and reverse class probabilites is discussed. For identification of ancient coins preselection on shape features and classification based on fusion of shape and a local features based representations is demonstrated.

4.1 Classification in Eigenspace

The conditional probability $P_s(\Omega_s | G)$ for an observed weight vector in the Eigenspace
Vector $\Omega_s$ depending on coin class hypothesis $G_s$ on side $s$ is estimated to be inversely proportional to the distance $D_s$ in the Eigenspace between the observed coin and a reference coin for the considered coin class.

A-priori probabilities $P_s(G)$ for coin classes on coin could be set to equal probability, if the coins are imaged in way that there is no rotation between obverse and reverse images, one can make use of this knowledge as modern coins are characterized by either 0 or 180 degrees of rotation between sides. In this case, the $P_s(G)$ are derived from the difference in rotation angles for obverse and reverse sides.

The fusion of probabilities for each coin side with prior information uses the Bayes formula:

$$P_s(G_s \mid \Omega_s) = P_s(\Omega_s \mid G_s) P_s(G_s),$$

where the denominator was omitted, as it is a constant term. Combination of both sides is done by the product rule (Kittler et al. 1998)

$$P(G \mid \Omega) = P(G=G_1=G_2 \mid \Omega_1, \Omega_2) = P_1(G_1 \mid \Omega_1) \cdot P_2(G_2 \mid \Omega_2).$$

4.2 Identification from Shape and Local Features

Local features based approaches and shape descriptors deliver distance measures between the coin in question and all other images in the database. In this case, a two-stage rank based strategy is possible, i.e. a small subset is preselected based on shape comparison and further processed using local features based matching (Huber-Mörk et al. 2008). Here, we follow a strategy combining probabilities which are derived from distance measures through a rule of combination (Huber-Mörk et al. 2010). Conditional independence between shape and local features, as well as between coin sides, can be assumed.

Ranking the shape dissimilarity $D_{AB}$ for shapes for shape $B$ matched to shape $A$ results in a preselection set $P$. From an observed shape description $A$ we derive a conditional probability for a coin side label $L$ assigned to $B$. The conditional probability for a $P_{\text{shape}}(L \mid A)$ is estimated to be inversely proportional to the dissimilarity between coins $A$ and $B$ labelled $L$:

$$P_{\text{shape}}(L \mid A) = 1 / \left( D_{AB} \sum_{C \in P} 1/D_{AC} \right),$$

where the summation term in the denominator accounts for normalization. A similar argument is applied to derive a conditional probability for observed local descriptors $X$ matched to local descriptors $Y$ labeled $L$ and corresponding to an image contained in the preselection set $P$:

$$P_{\text{local}}(L \mid X) = M_{XY} / \sum_{Z \in P} M_{XZ},$$

where $M_{XY}$ denotes the number of matches between the query image with local descriptors $X$ and the coin side image with local descriptors $Y$. The denominator accounts for normalization.

As local and shape features describe different properties of a coin, it is reasonable to assume statistical independence between shape and local feature measurements. Thus, the combination is performed by the product rule Kittler et al. (1998):

$$P(L \mid A, X) = P(L_{\text{shape}} = L_{\text{local}} \mid A, X) = P_{\text{shape}}(L \mid A) \cdot P_{\text{local}}(L \mid X),$$

where $L_{\text{shape}}$ and $L_{\text{local}}$ are labels derived from shape and local descriptions. The idea of fusion of different descriptor outputs is extended to a fusion of more than one image of a coin. Typically, a coin is presented by images of the obverse and reverse side. Fusion of coin sides is obtained in a straightforward fashion given by:

$$P(L \mid A_1, X_1) = P_{\text{shape}}(L \mid A_1) \cdot P_{\text{local}}(L \mid X_1) \cdot P_{\text{shape}}(L \mid A_2) \cdot P_{\text{local}}(L \mid X_2),$$

where $A_i$ and $X_i$ corresponds to shape and local feature descriptions of the $i$–th coin side.
5. Results

In this section we summarize results for classification of modern coins and identification of ancient coins. Figure 7 shows sample images for the considered collections of coins.

5.1 Classification of Modern Coins

Image data of modern coins was acquired through a coin collection which took place in the course of the implementation of the Euro currency in twelve European countries at the turn of the year 2001 to 2002. During this campaign 300 tons of coins coming from virtually all countries of the world but predominately from the twelve Euro member states have been collected and sorted by the Dagobert coin sorting system. Results are presented for a sample of 11,949 coins taken randomly from the collected money. A sample of coin images is shown in Figure 7(a).

A classification method should maximize correct classification for valid coins and correct rejection for invalid coins. Apart from FAR and FRR the case of wrong classification of a valid coin is also an undesired event termed false classification rate (FCR). FRR. From a receiver operator characteristics (ROC) curve, as shown in Fig. 8, the tradeoff between FCR plus FAR and FRR can be identified. An operating point, corresponding to a specific $t$, is found on the ROC curve, e.g. for perfect classification with FAR + FCR = 0, a very high FRR has to be taken into account (i.e. FRR > 0.5).

We discuss results including rejection based on the a-posteriori probability $P(G|Ω)$. A coin pattern $Ω$ is accepted to be of class $G$ if $P(G|Ω) ≥ t$, and rejected if $P(G|Ω) < t$, where $t$ is the rejection threshold. The parameter $t$ is used to tune the system towards the desired trade-off between false rejection and false acceptance.
Considering only valid coins, i.e. the 91.6% coins included in the 30 countries mentioned above, and using no rejection mechanism, correct classification was made for 98.27% of valid coins. With rejection aimed at minimization of false decisions a percentage of correct classification of 94.54%, 0.53% false classification and 4.93% false rejection is achieved for valid coins. Considering only invalid coins, i.e. the 8.4% unknown coins, classification into any of the known coin classes happens for 20.47%. Correct rejection of unknown coins is performed for 79.53% of invalid coins. Examining the mixed sample, a correct decision, i.e. correct classification or rejection, was made for 93.23% of all coins. False decisions, i.e. either false classification, false rejection or false acceptance, took place for 6.77% of all coins. Table 1 summarizes the final results.

A method based on direct matching of coin edge images was applied to a derived data set in an earlier study (Nölle 2003). The reported rather high false rejection rate of 14.22% was the main reason for the overall low rate of correct decisions of 85.58%.

### 5.2 Identification of Ancient Coins

We consider an ancient coin image database provided by the Fitzwilliam Museum, Cambridge, UK, which consists of 2400 images of 240 different ancient coins of the same class. Figure 7 (b) shows four of the coins contained in the data set. Each row shows the same coin acquired by different devices at varying conditions and different orientations. In particular, each coin side was acquired at three different angles of rotation using a scanner device and two acquisitions were made using a digital camera and varying illumination. At first sight, all coins bear the same characteristics. However, the coins shown in the different rows are produced by different dies. What makes this data set special and ideal to thoroughly test identification methods, is that all the coins are very similar. All the images are issued in the time of, or at least in the name of, Alexander the Great who came to power in Macedonia in 336 BCE and died as emperor in 323 BCE. Some of the coins are from much later and were minted in places around the Black Sea, in Egypt, in modern-day Turkey, Iran, etc. All coins follow the same basic standard: on the obverse side there is the head of Heracles in a lion-skin. The reverse side shows the god Zeus, seated left on a throne. Nevertheless, there is a huge range of detail in the minor variations that experts use to deduce the mint and date of the coin.

Shape matching is suited as a preselection step to the computationally less efficient matching based on local features which typically takes two orders of magnitude longer (Bay et al. 2006). The size of the preselection set is determined experimentally from Precision-Recall curves. Recall measures the ratio given by true positives divided by the sum of true positives and false negatives, i.e. $\text{rec} = \frac{TP}{TP + FN}$ and precision is given by $\text{prec} = \frac{TP}{TP + FP}$, where $FP$ is the number of false acceptance.

![Figure 9. Precision and recall versus preselection size for the ancient coins experiment.](image-url)
positives. Figure 8 shows plots of precision versus recall for the test set of 240 different images containing 10 images of each coin. Different settings of the shape matching weight parameter $\alpha$ show that a relatively large value of $\alpha$, which directs the matching dissimilarity towards more local influence, performs best. In order to obtain a preselection set of moderate size and high quality, i.e. the coin in question should likely be contained, a high recall is aspired. Figure 8 shows that a preselection set size of 9 to 10 images is suited for the task at hand.

For the experiments presented here, $P_{\text{shape}}(L|A)$ is computed for the 10 images with lowest dissimilarity and $P_{\text{local}}(L|X)$ for the same 10 images. The final decision is made according to the product rule. Table 2 shows the identification rates for the single descriptors and their combination with a leave-one-out evaluation scheme. The shape-based preselection of size 10 was performed accordingly to the given side of the test coin image. The DCSM alone gives an identification rate of 97.04% on the whole data set of 2400 images. For a preselection size of 10, there are only 13 cases (0.54%) where the correct coin is not contained in the preselected set. Consequently, local feature matching on the preselected set and fusion with the label probabilities from DCSM lead to an identification rate of 98.54%.

In an earlier study (Huber-Mörk et al. 2011), we compared a previous version of our shape matching approach to the shape context method (Belongie et al. 2002), and obtained 93.75% percent bull’s eye-test accuracy, whereas the state-of-the-art method obtained 50.64% percent on coin images.

### Table 2. Identification results for ancient coins using shape and local features matching.

<table>
<thead>
<tr>
<th></th>
<th>DCSM</th>
<th>SIFT</th>
<th>DCSM+SIFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>97.04%</td>
<td>71.77%</td>
<td>98.54%</td>
</tr>
</tbody>
</table>

6. Conclusions

We have presented methods for coin classification and identification applicable to coin collections comprising either a large number of coin classes, e.g. modern coins, or high intra-class variation, e.g. ancient coins. Modern coins represent financial value only if the coins are sorted and returned to the respective national banks. A tunable system is required as national banks accept coins only if they are delivered with a high degree of purity. The rejection mechanism based on the probabilistic fusion result allows to adjust a tradeoff between rigorous classification versus tolerant classification. Coin class probabilities for both coin sides are combined through Bayesian fusion including a rejection mechanism. Correct decision into one of the 932 different coin classes and the rejection class, i.e. correct classification or rejection, was achieved for 93.23% of coins in a test sample containing 11.949 coins. False decisions, i.e. either false classification, false rejection or false acceptance, were obtained for 6.77% of the test coins.

In order to facilitate prevention and repression of illicit trade of stolen ancient coins technologies aimed at allowing permanent identification and traceability of coins become of interest. Since every individual coin has signs, caused by minting techniques for pre-industrial ones or by use-wear for more recent ones, that make it unique and recognizable to an expert’s eye, traceability of pre-industrial coins can make use of visual inspection. Shape matching was used to match coin edges whereas the die of the coin was matched by means of local features. On a data set of 2400 coin images the combination of shape and local features outperform the accuracy rate of the single features and achieved an identification rate of 98.83%.

The results for classification of modern coins and identification of ancient coins are regarded to be almost perfect. Due to
large intra-class variance, the classification of ancient coins is still a challenging task, especially if attempted from single 2D images. Additional sensor information, e.g. from 3D image based measurements such stereo vision or photometric stereo or sensors providing information on material properties, could be used to overcome the difficulties inherent to 2D imaging of 3D objects. In a setup where complementary information is available, e.g. metadata or textual descriptions, the quality of automatic approaches to classification of ancient coins should improve significantly.

References


Archiving Three-Dimensional Archaeology: New Technologies, New Solutions?

Kieron Niven, Stuart Jeffrey and Julian D. Richards
Archaeology Data Service, University of York

Abstract:
This paper aims to provide a detailed overview of the issues involved in the creation, ingest, preservation and dissemination of 3D datasets from the perspective of a digital archive with particular emphasis on the topics of file size and format, metadata, documentation, dissemination and standards. The paper incorporates specific examples from past and present Archaeology Data Service (ADS) projects and highlights the recent work undertaken by the ADS and partners to specify standards and workflows in order to aid the preservation and reuse of 3D datasets.

Keywords:
Laser Scanning, Close Range Photogrammetry, Guidelines, Good Practice, Archiving

The huge growth of 3D data use in archaeology over the last ten years can be seen not only in the increasing availability of services and technologies that allow the collection of such data but also in the way in which such datasets often play a key or uniting role in larger, more diverse projects. The generation of 3D data occurs at numerous different scales, from landscape or seabed survey, through the laser scanning or photogrammetric survey of buildings and monuments, all the way down to the digitisation of small objects.

With such a pervasive and important role the issue of preserving such data for future re-use and re-interpretation arises. This is particularly relevant where the data is expensive to acquire or where it is used to monitor or ‘digitally preserve’ sites or objects that are either inaccessible or subject to deterioration.

The ADS has long recognised the issues associated with the creation, storage, preservation and dissemination of large 3D datasets and attempts to address the problems associated with such datasets can be seen in a number of past and current projects such as the ‘Big Data’ project (2006), VENUS (2009), the new Guides to Good Practice (2009-2010) and ADS current involvement in CARARE, investigating 3D object dissemination via European projects (http://archaeologydataservice.ac.uk/research/bigData; http://archaeologydataservice.ac.uk/research/venus; http://www.carare.eu/).

1. The Archaeology Data Service

The ADS was established within the University of York in 1996 as one of the five subject centres of the Arts and Humanities Data Service. Funding for the ADS came initially from the UK Arts and Humanities Research Board (now AHRC) together with the Joint Information Systems Committee (JISC) but currently consists of elements of core funding from the AHRC together with the Natural Environment Research Council (NERC) alongside a range of project-based funding from a variety of UK and European project and organisations.

The primary mission of the ADS is the long term preservation and dissemination of archaeological data via the internet. In support of this aim, the ADS also provides a technical advisory service to grant applicants and UK grant-giving agencies, develops and publishes a series of ‘Guides to Good Practice’, and makes available a range of teaching and learning resources.

Corresponding author: kieron.niven@york.ac.uk
A key component of these aspects of the ADS mission is the monitoring and recommendation of both file and metadata formats used in producing the many types of data created by archaeological projects. The data and metadata formats recommended in the ADS Guides to Good Practice have certain characteristics that enable them to address the numerous preservation issues highlighted in this paper and thus allow the data to be preserved and remain reusable for the long term.

2. Digital Preservation

“Digital information lasts forever - or five years, whichever comes first.” Jeff Rothenberg, RAND Corp., 1997.

The issues associated with the long-term preservation of digital data - together with the advantages of doing so - are becoming increasingly well-known across a wide range of fields. As a result, in recent years guidance and support has been developed at both national and international levels through a number of organisations and projects such as the Digital Preservation Coalition (DPC) and Digital Curation Centre (DCC) in the UK, the National Digital Information Infrastructure and Preservation Program (NDIIPP) in the US, and Digital Preservation Europe (DPE) and the Open Planets Foundation in Europe.

Within Archaeology, although awareness of the need to actively manage digital data is growing, practical developments towards actually doing so in a secure and standardised fashion are still some way behind other disciplines. As in the wider digital preservation sphere, the issues that are pertinent to the preservation of archaeological digital data revolve around the definition of standards and best practice i.e. what should be preserved and how this should best be done. A significant element of digital archiving focuses on the use and suitability of data file formats for the preservation and dissemination of data and involves such considerations as binary versus ASCII data types, proprietary versus open file formats, and the management of data compression. In addition, all data requires some form of documentation in order to be understood, not only in terms of how it came into being, but also what it represents and how it can be used. The specification of documentation and metadata standards, and their applicability to archaeological data, remains a significant digital preservation issue.

One of the most widely acknowledged approaches to the practical matter of preserving digital data for the long term is the Open Archival Information System (OAIS) reference model. OAIS comprises hundreds of pages of guidance and good practice and makes clear the importance of open file formats, data migration, robust and distributed hardware infrastructure and the necessity of discovery, access and delivery systems (Consultative Committee for Space Data Systems (CCSDS) Reference Model for an Open Archival Information System (OAIS) (CCSDS 2002). It does not, however, detail the specifics of the practical implementations of the processes it recommends. Actual digital preservation based on OAIS can be enormously complex. In archaeology preservation processes may have to deal with hundreds of file types, from hundreds of types of devices, using a plethora of software packages and the whole broad range of archaeological techniques. In addition, for a digital archive to be considered credible, and thereby attain ‘trusted digital repository’ status, it must be able to demonstrate well documented preservation policies and processes as well as having a robust long-term sustainability plan. As yet few formal mechanisms exist that act as a form of validation for digital repositories although one such is the Data Seal of Approval (DSA), an international peer review scheme (http://datasealofapproval.org/). The ADS was the second UK archive to receive the DSA after the UK Data Archive.
These concerns are equally present when dealing with the type of 3D datasets now frequently created from a range of techniques in many archaeological projects and often large 3D datasets can raise new or unique problems. The discussion of projects below aims to highlight the work the ADS has been involved in over the last six years in developing processes that address a number of these problems.

3. The Big Data Project

Many datasets generated by techniques such as marine bathymetry, laser scanning and LiDAR present specific challenges in that the volume of data created is frequently very large and therefore has storage implications (including the cost of buying and maintaining hardware or purchasing separate storage). In addition, beyond the actual storage of data, the physical size of such datasets can often create problems in terms of data access or reuse. The 2006 Big Data Project, funded by English Heritage, looked specifically at the practical issues raised in storing and disseminating large 3D datasets through three case studies: marine survey data from Wessex Archaeology, laser scanning data from Durham University, and LiDAR data from English Heritage (Austin and Mitcham 2007). The project started with a data audit and culminated with the deposit of data from each case study with the ADS. In addition the project also carried out a questionnaire survey and workshop aimed at ‘big data’ creators in order to quantify and assess the types of data being created alongside the options available for dissemination and reuse. As a result of these activities, the project produced a final report aimed at raising awareness of the issues associated with creating, storing, and accessing ‘big data’ as well as providing guidance in terms of both policy and practice.

The project report also provided a key set of recommendations for future research which has subsequently informed the recent Guides to Good Practice project (discussed below) with the findings incorporated into the relevant individual guides (Austin et al. 2008).

4. VENUS

In addition to the issues of storage and dissemination highlighted by the Big Data Project, ADS involvement in the 2006-9 VENUS project looked at the preservation of large, complex marine survey datasets, often featuring multiple streams of data combining to form various different data ‘products’ (Alcala et al. 2008). One key aspect of this project was to demonstrate how data selection plays a key role in producing a robust and reusable digital archive. The VENUS project itself aimed to develop scientific methodologies and deliver technological tools for the virtual exploration of deep underwater archaeology sites with the ADS role being focussed on the long term preservation of the project’s digital outputs. The ADS specifically focussed on the publication of a VENUS Guide to Good Practice alongside an exemplar digital archive (http://guides.archaeologydataservice.ac.uk/g2gp/ VENUS_Toc; doi:10.5284/1000004). A significant outcome of the project was the identification of ‘Preservation Intervention Points’ (PIPs) in the data lifecycle of the project. The VENUS underwater missions themselves surveyed shipwrecks at various depths by employing a complex data acquisition process using remotely operated unmanned vehicles with innovative sonar and photogrammetry equipment. Subsequent data processing stages also included the plotting of archaeological artefacts and the production of 3D models. At various stages in the data lifecycle, ADS identified PIPs where data was transformed by processes such as decimation, aggregation, recasting, or annotation, in addition to data being migrated from format to format. These stages were then evaluated in terms of whether, for the purposes of preservation, it might be appropriate to intervene and take a preservation copy of the data to be archived. The evaluation
process itself was based on a number of broad criteria, seven in total, which allowed each point to be weighed up in terms of categories such as reuse potential, repeatability, value (cost) and available metadata (in terms of both data reuse and the repeatability of specific processes). Interestingly the PIP criteria highlight that, although it is generally considered good practice to archive data in as raw a state as possible - because often the subsequent transformations applied can be recreated - this is not always the case for certain 3D datasets where data is subsequently merged (e.g. meshes) or processed (e.g. cleaned or decimated) to create ‘new’ interim datasets, often via a proprietary or automated processes.

As with the Big Data Project, the VENUS project has also made a significant contribution to the subsequent revision and expansion of the Guides to Good Practice through the further development of elements of the VENUS guide into a new general guide looking at marine survey data. In addition, the concept of PIPs is equally applicable to other datasets including laser scanning and photogrammetry and this conceptualisation of the process of data selection has also been incorporated into these guides.

5. Guides to Good Practice

The research that emerged from the Big Data and VENUS projects, as mentioned above, was given the opportunity to be incorporated into a wider suite of guidance material by the 2009-10 Guides to Good Practice project (http://guides.archaeologydataservice.ac.uk/) (Mitcham et al. 2010). The two-year collaborative project, funded by the US Mellon Foundation and English Heritage, aimed to revise, update and expand the original ADS series of Guides to Good Practice (six guides originally authored between 1998 and 2002 (Gillings and Wise 1998; Bewley et al. 1998; Richards and Robinson 2000; Schmidt 2001; Eiteljorg et al. 2002; Fernie and Richards 2002)) to cover a number of additional techniques including close-range photogrammetry (CRP), marine survey and laser scanning.

As with the previous series of Guides, the new updated series aims to specify standards and ‘good practice’ for a variety of techniques used within archaeological projects and covers the data formats that they produce, the suitability of these formats for long-term data preservation and, importantly, the metadata and documentation (at a number of levels) required to archive, understand and reuse these datasets in the long-term. Building on the work of the VENUS project, the new Guides also examine 3D techniques and data types such as laser scanning and photogrammetry within the larger project lifecycle of data acquisition, processing, reprocessing and the creation of various types of derived data such as CAD models, still images and video.

6. Current Projects

A number of recent and current projects are also contributing to the way in which the ADS approaches the archiving and dissemination of 3D datasets.

The deposit of data from the Virtual Amarna Project provided the first opportunity for the ADS to ingest and disseminate a large collection of 3D PDF files (doi:101.5284/1011330). A number of artefacts held in the site museum at Tell el-Amarna in Egypt had been scanned by a team from the University of Arkansas, working with Barry Kemp, the excavation director. As deposit of the archive coincided with the production of the Laser Scanning Guide to Good Practice, this provided the opportunity to create an excellent exemplar archive and highlight the various incarnations of data created within a laser scanning project together with the metadata required to understand and reuse this data (Limp et al. 2011; Payne 2011).

The ADS is also currently involved in the European CARARE project as a content
provider mapping and adding – amongst other data – 3D datasets to the Europeana portal. The use within the project of the 3D PDF format as the primary means of object dissemination has highlighted this format’s potential for easy and flexible dissemination of complex three-dimensional models as well as allowing ADS to work with similar European organisations that are creating or disseminating data in what is a relatively underused format.

In addition to 3D-specific projects, the ADS have also been undertaking work to ensure that the data it stores, regardless of type, is both secure and reusable. In 2011 the ADS was awarded the Data Seal of Approval verifying that it meets the standards of a trusted digital repository (Mitcham and Hardman 2011) In addition, in 2012 the ADS was also accredited as an official Data Archive Centre (DAC) for the UK Marine Environmental Data and Information Network (MEDIN) (http://www.oceannet.org/). The ADS have also recently implemented the ISO standard Digital Object Identifier (DOI) system for its collections via DataCite (http://datacite.org/). From an ADS perspective, the availability of DOIs for its digital collections not only ensures that data citations remain stable and resolvable but also that they are quantifiable in terms of reuse and impact via metrics from the DataCite DOI resolver.

7. Conclusions

3D data faces the same preservation issues as other archaeological datasets although, in some cases, certain problems are somewhat heightened, including issues of data storage, ensuring adequate metadata, documenting processing techniques, and dealing with proprietary software. While, for example, the use, and even preference, for proprietary or compressed 3D data formats within the community is a well recognised preservation issue, new complex problems in terms of access (due to the sheer size or number of files) is a particular characteristic of 3D data in archaeology. This increasing volume of data produced by 3D projects has another aspect in that, when the larger workflow is viewed and multiple incarnations of the ‘same’ data are seen as holding value, storing and providing access to these multiple sets of files becomes problematic. While storage capabilities continue to increase alongside decreasing hardware costs, indicating a possible solution to such data storage issues, it is worth noting that many technologies which capture 3D data continue to generate larger and larger volumes of data countering the perceived savings implied by lower hardware costs. It should also be noted that the most significant cost component of any digital repository is actually the labour required to ingest, migrate and subsequently manage the data. This cost is contingent on many factors, the least of which may be data volume (for example, see the breakdown of costs in the ADS Charging Policy - http://archaeologydataservice.ac.uk/ advice/chargingPolicy). The continuing growth and use of 3D data and the varied technical systems used to generate it will also hopefully continue to see a parallel development of both data format and metadata standards. Ongoing collaborative research at both a national and international level, as demonstrated by the projects discussed in this paper, will be an important factor in developing systems that ensure that the data produced remains secure and usable for the foreseeable future.

References


Intra-Site Analysis and Photogrammetry: the Case Study of the ‘Buca Di Spaccasasso’ (Grosseto, Italy) an Eneolithic Funerary Site

Giovanna Pizziolo, Daniele Pirisino, Carlo Tessaro and Nicoletta Volante
University of Siena, Italy

Abstract:
Here we present the outcome of a project that involved the implementation of an intra-site GIS based on photogrammetric methodologies at Spaccasasso Cave (Grosseto, Italy), an Eneolithic funerary site, the main burial chamber of which shows evidence of complex depositional practices. Due to their peculiar archaeological context, traditional excavation procedures are ineffective in recording these prehistoric palaeosurfaces. As a result, we have turned to using photogrammetric applications. The goal is to obtain 3D photorealistic restitutions in order to create DEM and ortophotographs, which are useful for managing archaeological information in a GIS. Two different methodologies are employed. The first is based on ERDAS IMAGINE LPS software, while the other makes use of open source software based on multi-vision scene reconstruction. 3D survey methodologies make it possible to render a realistic model of the archaeological site, which is an asset for recording qualitative data and 3D information, which is then used for performing intra-site analyses. We assess the different methods by focusing on the practical and logistical requirements of the excavation process.

Keywords:
Intra-site GIS, Photogrammetry, Copper Age

1. Introduction
This paper deals primarily with the archaeological application of digital photogrammetry and the implementation of GIS for taphonomic analysis. After a brief presentation of the archaeological context of Spaccasasso Cave, we focus on its morphological and archaeological characteristics, which have heavily influenced the choice of data-acquisition methodologies. Due to its peculiar archaeological context, traditional documentation maps are not effective tools for recording this prehistoric site; we therefore used photogrammetric elaboration. The paper presents the application of different methodologies based on photogrammetry and discusses the pros and cons of each application. The general assessment is intended to determine the most efficient tools for our real case study, not for an abstract or merely theoretical situation. We have tested whether the methodologies adopted fit well into the standard archaeological documentation procedure and allow us to carry out an intra-site analysis. The process of photogrammetric elaboration started in 2010, though only in an experimental phase, and in previous years was carried out via photographs which had not been taken with photogrammetric application in mind. The paper presents a comparison between the different methods adopted and provides an assessment focusing on the practical and logistic needs of the excavation process.

2. The Buca di Spaccassasso: Archaeological Background

The Buca di Spaccassasso Cave is a prehistoric site located on a calcareous hill in Maremma Regional Natural Park in Uccellina (Southern Tuscany, Grosseto, Italy). The cave is characterised by thermal action, which causes mercury sulphide to be deposited on the limestone rock and results in cinnabar veins. Cinnabar minerals were harvested in different phases of occupation, and the site shows evidence of occupation over a long period of time, which offers the main proof connecting
it to the Copper Age, the Early and Middle Bronze Age, and Late Antiquity (V-VI cen. A.D.). During the Late Copper Age the site was also used as a funerary area, characterised by secondary deposition rituals.

The archaeological investigation, which is still in progress, was begun by the Board of Archaeological Heritage in Tuscany (Soprintendenza per i Beni Archeologici della Toscana) in 2000, and has been overseen by Siena University since 2004. The excavation is focused on the funerary area located in the external side of the cave and at the base of a sub-vertical rock shelter (Fig. 1). The area is delimited on the western side by three boulders and on the southern and northern sides by medium-sized stones that form a sort of wall. These features border the funerary chamber, which is about 2.5 meters wide. Due to its particular natural location and the style of its pottery, Spaccasasso can be classed as part of the ‘Grosseto Group’ (Leonini and Volante 2005). The stratigraphy shows evidence of complex secondary funerary rituals, in which substantial deposition occurred between the Late Copper Age and Early Bronze Age. Similar funerary rituals are not so common for the Late Eneolithic in central Italy, but are in fact more common in northern Italy. In secondary funerary rituals, human bones were laid down following the initial removal of flesh, sometimes resulting from the effects of a fire that took place somewhere else. At Spaccasasso, human bones and grave goods have consequently been displaced and arranged in the funerary chamber according to the different deposition processes, resulting in a chaotic array of bones. They are arranged out of anatomical order, are fragmented, and are mixed with other items such as pottery and to a lesser extent lithic artefacts and beads. In several cases, layers of stones were prepared to support or to cover unarticulated bones that had been either arranged randomly or spatially organised according to different conditions.

3. General Aims and On-Site Methodology

Due to this complex situation, it is fundamental that taphonomic analyses are carried out very carefully, using the maximum amount of detailed information acquired during the excavation, and with the additional aim of using the information related to the depth of bones and palaeosurfaces. For this reason, the standard way of drawing maps on site is insufficient for recording all the data related to bones, artefacts, and possible stone floors without the documentation process being very time-consuming and affecting the logistics of the excavation. Therefore, since 2005 some experimental strategies have been applied using different methods of photographic documentation. Since then we have chosen the following strategy: after a quick rectification is carried out, photographs are printed on-site so as to use them to record data during the excavation. The photo prints are then used as a ‘background’ onto which the ID and x-y-z coordinates of the centre of every single bone, artefacts and stone (when necessary) are recorded, while every shape is outlined in different colours according to their characteristics. These sketches represent raw information, which is then developed by means of the methods that we will discuss in the following paragraphs to obtain regular archaeological documentation maps.
4. Workflow and Methodologies

Once back at the university laboratory, the photos and the information recorded on them are processed in different ways according to the different methods adopted.

This process has been developed by creating a GIS which works as a general archive for storing raster and DEM data obtained with photogrammetry. The GIS has made it possible to create vector information layers to be used in archaeological maps (Fig. 2). The archaeological features are organised in the GIS (ArcGIS Esri®™) system within a geodatabase, in which classes of artefacts are drawn as polygons. These classes are described according to their characteristics, which, in the near future, can be updated further through specific studies. Moreover, every stone is entered into the geodatabase according to its lithology, while the stratigraphic units are characterised by taking into account the qualitative information either recorded on site or visible in the photographs. In this way, the GIS stores information obtained from firsthand observation recorded during the excavation (such as traditional maps, sketches, forms, etc.) as well as indirect information obtained through the analysis of photographic documentation. The structure is conceived as an open system in which the metadata of every feature makes it possible to compare data of varying types. This comparison is an efficient way of cross-checking the overall qualitative and quantitative information. In addition, the GIS system collects other information related to the Spaccasasso site, and it provides the most suitable tools for performing a taphonomic analysis by means of thematic maps and spatial analysis.

It is clear that a simple rectification of photographs, shot in a sub-zenithal way on the excavation, is not sufficiently spatially accurate to be used for drawing archaeological maps. In fact, apart from camera lens distortion, geometric problems arise with regard to the three-dimensional effect. In order to tackle this problem, we have tested two photogrammetric approaches with the aim of comparing the pros and cons of different methodologies. The goal is to test an effective way of developing an acceptable workflow to record taphonomic processes in the Spaccasasso excavation; that is to say, the aim is to document briefly a complex archaeological stratigraphy located on the rocky side of a calcareous hill. As a matter of fact, the physical characteristics of the area, which is very difficult to access, suggested the need to employ highly portable technologies and to facilitate the creation of reliable 3D models and orthophotos upon which to digitise elements of interest with the associated database. It was thus considered convenient to explore the capabilities offered by digital photogrammetry in order to evaluate the possibility of a more systematic use of reality-based modelling at Spaccassasso. The photogrammetric technique

\[\text{Figure 2. A layout of one archaeological stratigraphic unit of the Spaccasasso Cave site. The map shows orthorectified photos on which the archaeological information organized in a Geodatabase are overlaid. The orthophotos have been used to acquire the shape of archaeological artefacts.}\]
provides new data for our taphonomic studies and offers the invaluable benefit of being able to recover 3D data that would otherwise be lost. Part of our project is dedicated to the use of photogrammetry and to the use of photographic documentation taken years ago when we did not foresee any specific use for photogrammetric elaboration.

Before attempting photogrammetric elaboration, we have experimented with the use of laser scanner techniques on the Spaccasasso site. The scanning, carried out by Tesette with a Laser Scanner (Mensi GF200), was executed with TOF technology (optimal range 2-100m distance with millimetric accuracy), obtaining different scanned areas, which have subsequently been merged. The scanning was set up every 3mm along each surface, and the laser was placed in different positions in order to scan the funerary chamber, the rock shelter, and the cave.

The data have been used to reconstruct the whole funerary environment and to explore the spatial relationship between these features in order to understand the ‘scenographic’ settings of the funerary area. Moreover, the scanning of the funerary chamber was used as a reference measurement in order to compare it with other 3D reconstructions of the palaeosurfaces. Evidently the scanning techniques represent the best solution for obtaining 3D data from a surface; however, in practical terms it is not feasible to use this tool continually on the excavation to document the different archaeological layers. In this archaeological context, we need 3D data for every single archaeological layer in order to develop a taphonomic analysis, and we need to link that information to our geodatabase.

5. First Use of Photogrammetry: Leica Photogrammetry Suite

The decision to resort to photogrammetric methods was made in the fall of 2010, after the conclusion of that year’s campaign. At this point, it is relevant to discuss the experimental nature of the way in which the project was first carried out. Indeed, our workflow was heavily affected by the fact that the set of data available had not been recorded with the use of digital photogrammetry in mind, and that not all the imagery at our disposal could successfully be used in the photogrammetric process. Notwithstanding that, it has been possible to generate very accurate DEMs and orthophotos for the implementation of an intra-site 3D GIS. In addition, the project tested the validity of the excavation records and highlighted the interaction between the data produced with traditional archaeological documentation techniques (both written and visual records) and the operational procedures of the software available to us, namely Leica Photogrammetry Suite 9.2.

The first step required a thorough examination of the excavation’s archive in order to collect the necessary data, such as digital images and ground measurements. The most suitable photographs were selected from among the many available according to the criteria of area coverage and image overlap. Special attention was paid to the uniformity of the principal factors that affect the scale of the images, namely the focal length and the camera’s height above the ground. The consistency of camera rotation angles were also taken into account, and the best results were obtained by processing images produced with a minimal angle of inclination in relation to the vertical. However, the absolute validity of the imagery was assessed following a somewhat time-consuming and empirical process, and several projects were run with different sets of data. Most of the photographs available were produced so as to represent one square of the excavation grid at a time; however, our archaeological focus did not necessarily revolve around contiguous areas of the grid, and the

2 The procedures of use for the software are briefly summarised in the present paragraph. An exhaustive analysis of the settings and operational workflow can be found in the user guides 2008 ‘Project Manager’ and ‘Automatic Terrain Extraction’ of Leica Photogrammetry Suite software.
sequence of the exposure stations did not follow the logic of a continuous strip or a consistent direction of flight, nor were the neighbouring grid squares always photographed at the same stage of the excavation. Most of the overlapping photographs thus displayed only one square at a time, with a very high degree of overlap percentage (often exceeding 85%). There were also gaps of unrecorded space between adjacent excavation grid squares which were not covered by the overlapping images. For this reason, modules of two photographs had to be used for each one of the required photogrammetric restitutions, as it has not been possible to process blocks containing larger numbers of images.

Another essential issue to overcome was identifying and retrieving the ground control points coordinates for the successive steps of the photogrammetric workflow. These were recovered from the set of excavation plans, which were traced in the field using the photo print-outs as a reference. The excavation plans display many of the features that are also identifiable in the images used for the photogrammetric process, and clearly indicate the measured elements of interest (e.g. artefacts, bones, removed stones, grid vertices). Each element of interest was assigned an identification number and was recorded on catalogue sheets along with all the necessary information, including their spatial coordinates, since they were measured on the field by means of optical instruments. In this way, a redundant number of ground coordinates were collected and used as ground control points. Moreover, some of the measured features were used as checkpoints to assess the overall accuracy of the triangulation solution. The tie points, on the other hand, were automatically generated by the software. Overall, the retrieval of the data through the examination of the excavation archive proved effective. In fact, the outcomes of the aerial triangulations of the most successful projects yielded a very high standard of accuracy, and the residuals of both control points and checkpoints were kept below a very low threshold (Fig. 3).

Having obtained good triangulation results, it was possible to automatically generate high quality elevation models. Extraction reports and Point Status images were used to verify the validity and accuracy of the output DTM.s. The following stage was the creation of orthorectified images, for which the DEMs were used as elevation source in the ortho resampling process. Afterwards, the 3D digital representations and the related orthophotos were uploaded into ArcGis, where the excavation view could be reconstructed, and these data were used as a background reference for the digitisation of the elements of interest and the implementation of the excavation GIS.

To conclude this concise description, the results of the photogrammetric processing and
the 3D GIS implementation were satisfactory. However, having to deal with incomplete and sometimes unsuitable data proved to be complex and time-consuming. In fact, the results of this work again revealed the importance of the most significant phase of the whole photogrammetric pipeline, namely project planning and the acquisition of the images. On the other hand, it has been possible to fully exploit the data produced in the field and to appreciate the potential of traditional archaeological documentation, even when its recording strategy has not been specifically devised for digital photogrammetric applications.

Nevertheless, our experiment was taken further, and the following paragraphs will discuss the employment of computer vision techniques in order to process the data from the subsequent campaign of excavations. In fact, given the validity of the 2010 results, we decided to use the Agisoft PhotoScan software to process the new 2011 set of data.

6. Computer Vision

Our first attempt aimed to increase the informative content of our existing 2D documentation, in order to render it into 3D documentation. The application of computer vision techniques can achieve this goal with an easy, fast, and low-cost solution. As far as the management of the data in the laboratory and during the excavation is concerned, this method is very convenient. In the field, only the standard equipment needed for the survey is necessary, such as a total station for measuring certain ground control points (which is useful for georeferencing the 3D models) and a standard digital camera for taking a sequence of images, which is the only necessary tool for the photogrammetric 3D reconstruction.

Computer vision (CV), sometimes called machine vision is a field of computer science that includes methods for the reconstruction and description of 3D reality from 2D images. It is a scientific discipline that reconstructs, with mathematical methods, the 3D morphologies of objects represented in the images (Szelisky 2010) (Lillesand, Kiefer and Chipman 2007):

The goal of the computer vision is the construction of scene description from images [...] (Shapiro-Stockman 2011).

State-of-the-art photogrammetric methodologies related to CV can provide archaeology with a methodology of surveying that is easy to manage, completely automated, and low-cost in comparison with other methods of 3D reconstruction, such as laser-scanning techniques.

There are a variety of different tools that have been developed as a result of CV and which are useful for 3D reconstructions based on 2D data. One such example is that of multi-view stereo algorithms. They differ from the stereo algorithms that produce dense depth maps using a pair of binocular images, in that their goal is to reconstruct a 3D scene from a set of images taken from known camera viewpoints. 3

Our work is not the first example to employ imaged-based 3D modelling in an archaeological context (Anderson 2010; Campana and Remondino 2008; El-Hakim et al. 2008; Ducke et al. 2011; Verhoeven 2011; Doneus et al. 2011), but it seems to stand out due to its development of a concrete surveying method based on a 3D photogrammetric reconstruction, and its goal to improve the standard of quality of the documentation and to increase the possibility of providing a better interpretation of the archaeological context.

The methodology can be summarised by the following set of tasks, and the workflow is an integration of different software:

- Unordered image sequence
- Structure from motion

3 A quantitative comparison of several multi-view stereo reconstruction algorithms are available online at http://vision.middlebury.edu/mview (Seitz et al. 2006).
Intra-Site Analysis and Photogrammetry: the Case Study of the ‘Buca Di Spaccasasso’

Giovanna Pizziolo et al.

Figure 4. Spars 3d point cloud obtained from Bundler and managed and visualized in Photoscan. The image shows the 3d information and the relative positions of the camera when the photos have been shot.

Figure 5. Dense point cloud.

Figure 6. Surface reconstruction on which the orthophoto has been draped.

- Feature extraction
- Image matching and camera reconstruction (3D sparse point cloud) (Fig. 4)
- Dense 3D point cloud (Fig. 5)
- Surface reconstruction (Fig. 6)
- Generation of a georeferenced 3D model and processing of DEM (digital elevation model) and ortho-photos.

This method consists of the integration of different software. The first step is achieved with the use of the open source software Bundler (http://phototour.cs.washington.edu/bundler/) (Snavely, Seitz and Szeliski 2007). This first step can be broken down into smaller sequences: feature extraction (keypoints), image matching, camera position, and sparse 3D point reconstruction.

The process starts with an unordered sequence of digital images. The next step is features extraction, consisting of an automatic identification of characteristic regions for every image, and the task is further expanded by using the Scale Invariant Feature Transformation (SIFT) algorithm (Lowe 2004). The software implementation of SIFT made by its author allows the extraction of around 2000 features for every image of 500 x 500 pixels (Lowe 2004). In this case, the features are the regions where a particular variance of grey values are recorded, since the software implementation of the algorithm works with the copies of grey scales of the original images. The task consists of using this relation data in combination with the focal length and camera CCD width, which are parameters recorded inside the exif properties of the files. A technique known as bundle adjustment is adopted for reconstructing 3D scene geometry. The Bundler software is a structure-from-motion software that works in this way. This technique does not require any kind of parameters such as, for example, the specification of the spatial coordinates of the camera, or a set of ground control points for image rectification. The result of the Bundler elaboration is a 3D sparse scene geometry.

This output can be used to generate a 3D dense point reconstruction. There are different approaches for generating a 3D dense point cloud. One of the most accurate algorithms is PMVS (Patch-based Multi-View Stereo) (Furokawa and Ponce 2007). In any case, there are various pieces of software dedicated to this technique, some of which are open-

---

4 The software’s application of SIFT algorithm was written the first time by its creator D. Lowe, but it has not been released in an open license. It is, however, possible to use it for educational and research purposes: US Patent 6,711,293 (March 23, 2004). The software link is available at http://www.cs.ubc.ca/~lowe/keypoints/.
### Table 1. Some processing parameters and statistics derived from the z values inside the square selected area.

<table>
<thead>
<tr>
<th>Method</th>
<th>Time (s)</th>
<th>Points number</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Dev. st</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMVS</td>
<td>960</td>
<td>27239</td>
<td>115.1450</td>
<td>115.3593</td>
<td>115.2192</td>
<td>0.04461</td>
</tr>
<tr>
<td>Photoscan (point cloud)</td>
<td>12</td>
<td>3540</td>
<td>115.1692</td>
<td>115.3239</td>
<td>115.2230</td>
<td>0.031</td>
</tr>
<tr>
<td>Photoscan (lowest)</td>
<td>3782</td>
<td>3560</td>
<td>115.1611</td>
<td>115.3415</td>
<td>115.2279</td>
<td>0.0365</td>
</tr>
<tr>
<td>Photoscan (low)</td>
<td>454</td>
<td>14161</td>
<td>115.1516</td>
<td>115.3472</td>
<td>115.2267</td>
<td>0.03819</td>
</tr>
<tr>
<td>Photoscan (medium)</td>
<td>3019</td>
<td>57364</td>
<td>115.1476</td>
<td>115.3593</td>
<td>115.2222</td>
<td>0.04072</td>
</tr>
<tr>
<td>Laser</td>
<td>-</td>
<td>5034</td>
<td>115.1586</td>
<td>115.3695</td>
<td>115.2275</td>
<td>0.045</td>
</tr>
</tbody>
</table>

source and which may be based on server webs such as Arc3d (http://homes.esat.kuleuven.be/~visit3d/webservice/v2/) or stand-alone platforms such as Photomodeler (http://www.photomodeler.com/). It is important to underline the hypothetical accuracy of the result: more than 1/200 (1mm for an object of 20cm) from a sequence of images with a resolution of less of 640 x 800 pixels (Furokawa and Ponce 2009).

The last step consists of the formation of a three-dimensional surface defined by the interpolation of the 3D dense point clouds. Some particular interpolation techniques are necessary for defining the surface (a mesh) that can be easily managed with the OS software Meshlab (http://meshlab.sourceforge.net/). The most popular algorithms are ball pivoting (Bernardini et al. 1999), poisson reconstruction (Kazhdan, Bolitho and Hoppe 2006), power crust (Amenta, Choi and Kolluri 2000), and tight cocoon (Dey and Goswami 2003). The specific method best applied is determined by the particular context. Poisson reconstruction appears to be a universal technique and it generally provides good results in all cases (Ducke 2011). Ball pivoting gives an accurate surface reconstruction, which saves time, but tends to leave many holes in the surface.

At the end of the process we need to produce a 3D georeferenced model. This step requires a complete data set of digital elevation models of units of stratifications (Harris 1975). For this task we used the software Photoscan, produced by Agisoft (LCC 2011). This software offers a complete suite for the creation of 3D models, including the integration of SFM techniques. Photoscan is also capable of managing the 3D meshes and textures that allow for the possibility of importing the result of PMV/CMVS processes through the use of a PLY file extension. Before the photographic survey can be carried out, it is necessary to define some ground control points, which are useful for georeferencing the 3D model with Photoscan and for managing the surface in a georeferenced coordinate system. This is used to export the results with file extensions that are capable of being fed into a GIS, such as DEM (ASCII-grid) and orthophoto-mosaics (.jpeg or .tiff). These models are fed into a geodatabase where the archaeological documentation is stored. Inside this platform, the raster and vector datasets are correlated and table archives are derived from the archaeological surveying on the excavation.

### 7. Comparison of Methods

It is very easy to manage and create 3D models with the computer vision technique. However, it is necessary to test the method’s accuracy before confirming its use as a standard method of mapping and documenting an archaeological excavation. In order to assess the final accuracy of different modelling techniques, we should compare the results achieved with the laser scanner and the CV 3D model.

In order to assess the method’s accuracy, an area of 50cm² was selected, which was located in the middle of the burial chamber. This area is characterised by a very non-homogenous...
Intra-Site Analysis and Photogrammetry: the Case Study of the ‘Buca Di Spaccasasso’
Giovanna Pizziolo et al.

Figure 7. DEM surfaces obtained from laser-scanner (1), Photoscan (medium) (2), PMVS (3) and the position of two profiles A-B.

The 3D CV reconstruction was accomplished using Photoscan for the georeferencing process and for exporting the results into grid ASCII files. After this step, the grid ASCII files obtained with Photoscan and the laser-scanner were entered into ESRI’s ArcGis 10 as a point geometry shapefile. These feature classes were then interpolated with the same method (ordinary kriging) and parameters, with a resolution of 0.003m cells (Fig. 4). This dimension is, to a certain extent, similar to the grid density of the laser scanner.

The laser-scanner result was considered the basis of comparison, and was treated as if it were a 100% accurate result. By comparing the difference between the CV and laser-scanner elaborations (Fig. 7), several metric and statistic values are extracted from the georeferenced surfaces (Table 2).

The amount of variation between the digital elevation models can be measured using a standard deviation. In this case the values show a similar distribution between the Photoscan medium and the PMVS algorithm. The RMSE (root mean square error) was used with the aim of illustrating the vertical position accuracy. It is similar to a standard deviation but is based on the assumption that errors in DEM are random and normally distributed (Wechsler, 1999). In this context, the accuracy is expressed in ground distance at a 95% confidence level. The reported accuracy value reflects, in theory, all uncertainties, including those introduced by geodetic control coordinates, compilation, and final computation of ground coordinate values in the product. The RMSE values show that the most accurate surface, as far as the vertical accuracy is concerned, is obtained using PMVS (0.038). Nevertheless, all other results are not greatly inaccurate in terms of the vertical metric accuracy.

Moreover, an analysis of the raster maps (obtained using the difference between the DEM derived from the laser-scanner and all other 3D surfaces) shows how the PMVS algorithm tends to delineate a more homogenous surface (Fig. 6).
Method | Min (diff) | Max (diff) | St. Dev | Mean | RMSE*  
--- | --- | --- | --- | --- | ---  
Photoscan (point-cloud) | -0.075 | 0.065 | 0.0189 | 0.00438 | 0.044  
Photoscan (lowest) | -0.096 | 0.051 | 0.018 | 0.00057 | 0.037  
Photoscan (low) | -0.092 | 0.061 | 0.019 | 0.00073 | 0.040  
Photoscan (med) | -0.92 | 0.061 | 0.019 | 0.00053 | 0.043  
PMVS | -0.97 | | | |  

Table 2. Statistics and RMSE* (Root-Medium-Standard-Error). The values are the result of the difference between the laser scanner and the other 3D surfaces.
Intra-Site Analysis and Photogrammetry: the Case Study of the ‘Buca Di Spaccasasso’
Giovanna Pizziolo et al.

Time in data acquisition/elaboration | Handy in field | Re-use recycle data | Output
--- | --- | --- | ---
Laser Scanner | One day | Not much portable | Not possible | DEM 3D model

Leica Photogrammetry Suite

One day | Yes but need of “ad hoc” set of photos, difficult to shot | Possible but with some limits in making use of a large number of images together | DEM orthophoto

Computer Vision

3 hours | Yes, not specific requirements for taking picture | Good flexibility in using “old” photos. High level of re-use of previous information | DEM Orthophoto 3D model

Table 3. Assessment of different methods.

(50 cm² areas) which has been georeferenced in the UTM ED 50 reference system.

At Spaccasasso, we have applied different survey techniques which showed the pros and cons regarding their flexibility, according to the needs of our archaeological documentation. The assessment of the different methods is summarised in the following table (Table 3). To sum up, we can affirm that photogrammetry proved to be the most appropriate solution for dealing with the problems of the archaeological archive of Spaccasasso. Our need for spatial accuracy was fulfilled through all the methods implemented. However, there are some differences concerning the practical side of the excavation process, as well as the amount of time required in the laboratory to manage the 3D elaboration and its GIS implementation. Moreover, it is important to note that CV allows for an easier treatment of ‘old’ photographs, i.e., those related to excavation layers shot in previous years and which had not been taken with photogrammetric application in mind. These photos, in fact, do not follow any specific schema and are obtained from different points of capture. The re-use of old photographs is a great advantage in the management of the whole archaeological recording process of the site.

Indeed, we can conclude that CV gives the best results in terms of photograph management and 3D data processing time.

8. Conclusions

In conclusion, computer vision could make it possible for GIS to become a standard for 3D archaeological documentation. CV can re-render 2D data into a complete 3D reconstruction of stratigraphy by means of spatial (geometrical) models for GIS implementation, such as 3D grid (Voxels) (Bezzi et al. 2006) or 3D boundary representation (Multipatch Geometry Type – Esri®)(Stoter and Zlatova, 2003; Katsianis et al. 2008; Lieberwirth, 2007).

A complete 3D restitution can be employed for two main purposes. The first is to create a 3D model of the archaeological stratigraphic deposit; the second is to use the model as a reference to perform spatial analyses in a volumetric space and to explore the spatial relationship between the artefacts and the stratigraphic units. Excavation and study are still in progress at Spaccasasso Cave: thus we will discuss the results of the whole system of documentation and elaboration of taphonomic data in future papers.

Acknowledgments

We are sincerely grateful to Paolo Machetti, Moris Faccone and Vincenzo De Troia of Tecsette srl for their availability to scan Spaccasasso Cave and to collaborate with us in the field.
This work has been carried out in full collaboration by the Authors. In particular paragraph 2 has to be attributed to N. Volante, paragraph 3 to G. Pizziolo and N. Volante, paragraph 4 has to be attributed to G. Pizziolo, paragraph 5 has to be attributed to D. Pirisino, paragraph 6 has to be attributed to C. Tessaro, paragraph 7 has to be attributed to G. Pizziolo and C. Tessaro.

References


Site Recording Using Automatic Image Based Three Dimensional Reconstruction Techniques

Victor Ferreira, Luís Mateus and José Aguiar
Universidade Técnica de Lisboa, Portugal

Abstract:

The objectives of this paper are to demonstrate that image based three dimensional reconstruction techniques, such as structure from motion (SFM) and multi-view-stereo (MVS) enable the production of documentation with comparable quality to the one produced with terrestrial laser scanning (TLS). We start by demonstrating that proposition and then, will present two more case studies where those techniques were used to produce data for the archaeological recording. In one of those, SFM/MVS was combined with TLS, providing a new layer of information after excavation was done.

Keywords:
Photogrammetry, Structure From Motion, Archaeology, Terrestrial Laser Scanning, Point Clouds

1. Introduction

Traditionally the archaeological recording is based on direct measurements made upon a grid set up on field. Used as a coordinate frame it supports manual drawing done on sheets of paper. The setting out of the grid could be accomplished with the use of a theodolite. Total stations allowed not to materialize a grid on field and to consider it only conceptually. Either way, the recording is always based on the discretization of the archaeological structure towards a two-dimensional representation. In recent times, the use of terrestrial laser scanning (TLS) resulted in a much faster and reliable way of recording with the advantage of three dimensional data acquisition (Borrazás et al 2008) (Mateus et al 2008) (Costantino 2010). Usually it is either used to generate colour orthoimages or to enable direct 3D drawing. Field recording is relatively straightforward but we think it is still an expensive technology requiring an high level of expertise.

The very recent developments of automated three dimensional recording techniques based in images brought a new set of opportunities for the archaeological recording that, in some circumstances, present advantages with respect to TLS. Examples of commercial software tools that implement this kind of techniques are Image Master (http://www.topconpositioning.com/products/software/office-applications/imagemaster), Photomodeler Scanner (http://www.photomodeler.com/products/pm-scanner.htm) and PhotoScan (http://www.agisoft.ru/products/photoscan). There are also free web services - My3DScanner (http://www.my3dscanner.com/), Microsoft Photosynth (http://photosynth.net/); Arc3D (http://www.arc3d.be/), Autodesk 123D Catch (http://www.123dapp.com/catch), and freeware software available on the internet such as VisualSFM (http://www.cs.washington.edu/homes/ccwu/vsfm), Photosynth toolkit (http://www.visual-experiments.com/2010/08/19/my-photosynth-toolkit/), and PMVS/CMVS (http://grail.cs.washington.edu/software/cmvs/). All these freeware tools are based in the structure-from-motion (SFM) approach for camera orientation and in the multi-view-stereo approach (MVS) for dense point cloud reconstruction.

In this article, we present three case studies in the “Convento de Cristo” UNESCO World Heritage site in Tomar, Portugal. It was used a combination of Visual SFM - Visual Structure From Motion (Wu 2007) (Wu 2011) and PMVS/CMVS - Clustering views for
Multi-View Stereo / Patch-based multi-view stereo software (Furukawa and Ponce 2009) (Furukawa et al. 2009). The reason for this choice, although it may not ensure the most accurate results, is due to free access to the software and with the possibility of controlling all the process since the processing is done locally. It was also important to assess to what extent these freeware tools can provide reliable results.

First we start by demonstrating that image based techniques, such as the selected ones, can provide data with acceptable metric quality when compared with TLS, using an example where the highly decorated “Janela Manuelina” of the “Convento de Cristo” was recorded.

Secondly we discuss the case of “Pátio dos Carrascos” where those techniques were used to retrieve data for the archaeological recording. We demonstrate that the semantic quality of the data was suited to the archaeological recording. By semantic quality we mean visual qualities of the obtained materials that can be used for the archaeological interpretation and analysis.

And thirdly we discuss the case of “Paços do Infante”, where these image based techniques allowed the generation of 3D data defining a new layer of information added to a previously survey done with TLS. This operation allowed showing that both surveys have comparable performances both in metric and semantic aspects. It further allowed to visually depict the amount of excavation that was done.

2. The SFM/MVS Workflow

In general, image based three dimensional reconstruction workflow can be divided in two stages: i) image acquisition, and ii) image processing.

The first stage has to follow some specific rules to enable a more successful processing which is almost unattended with the SFM/MVS approach. It is recommended that features to be reconstructed should appear in three or more images. Images have to be redundant, and consecutive images should be taken with small base distances. The total number of images can be several hundreds. The final spatial resolution of the point cloud models is directly related with the camera distance to the object and with image resolution.

The processing stage consists on the following steps: i) feature detection in images, ii) image matching, iii) camera calibration, iv) camera orientation (relative orientation), and v) geometry reconstruction. In the SFM/MVS approach, geometry reconstruction occurs in two steps: i) sparse reconstruction, and ii) dense reconstruction. Usually, dense reconstruction uses data retrieved in the previous step as input data. The user only has to select what images are going to be processed and to define a set of parameters that will have effect on the quality and density of the reconstruction. After the reconstruction is complete, it is possible to scale and orient the model for further processing (ortho image generation, mesh generation). Until recent times these steps were done manually (Mateus 2008) (Almagro 2008) (Drap 2009). It was a cumbersome and time consuming task where an operator had to pick multiple points on multiple images. Today it is possible to perform all those steps in a fully automatic way. We may coin this form of photogrammetry as automatic digital photogrammetry (ADP).

In our examples, for the sparse reconstruction we used the software VSFM and for the dense reconstruction we used the software CMVS/PMVS. This last tool uses the processing data produced in the first processing stage by VSFM, and has two steps: i) clustering of groups of images (CMVS) and ii) dense reconstruction of 3d points (PMVS).

CMVS analyzes the image data, and tries to divide it in groups that have highly
correlated images, and with some of them also in the other groups. The maximum size of the groups is selected by the user and is dependent of several factors: image size, available RAM (critical factor), processing time. Normally, in the first time a job is processed, the user should test the group size to make sure the processing does not exceed the available RAM. CMVS creates files with processing options for each group created, so PMVS (in the next stage), knows how to process them.

PMVS is invoked, processing each group sequentially. It is also possible to divide the groups by several computers shortening the total time of processing.

There are a set of parameters that have to be adjusted in order to control the reconstruction. Some relevant parameters are level, csize, minImageNum, CPU, quad, maxAngle. Basically, with these parameters it is possible to control the downsampling of the original images, the density of reconstructions, the minimum number of images where a point as to be visible to be reconstructed, the number of CPUs to be used, the noise of the reconstructions, and the angle between homologous rays. Full documentation about the meaning of these parameters and default values used can be found in PMVS home page (http://grail.cs.washington.edu/software/pmvs/documentation.html). When starting to use PMVS, one can use it with the default configuration (that is acceptable to the average object), or change the parameters, in order to get a better result, depending on the purpose of the final model.

This automatic approach is highly dependent on the texture of objects since the algorithms used rely on matching features identified in multiple images. So, features must exist, otherwise automatic reconstruction is not possible. Usually, historic buildings are rich in textures. This can result in very dense textured point clouds that in some cases can equal or supersede, in density, the TLS point clouds.

3. Case Studies

Here we describe the image acquisition strategy, the processing workflow, the results obtained and the comparisons with the other recording techniques referred (traditional techniques and TLS). In Fig. 1 the three recorded places are identified.

3.1. Assessing the Metric Quality of a SFM/MVS Point Cloud Model - “Janela Manuelina”

One of the more known places of the “Convento de Cristo” is the famous “Janela Manuelina”. It is a very complex and detailed piece of sculpture that presents a serious challenge of recording. It measures approximately 5mx12m. To be able to accomplish a full and comprehensive photographic recording we used a telescopic mast. This enabled us to get images from higher angles, so that blind spots could be minimized. We also defined an approximate half circular pattern to displace the camera ensuring that small base distances were kept. Compliance with this principle was assured both vertically and horizontally.

The images were then processed with VisualSFM software, making a sparse reconstruction and then using CMVS/PMVS software for the dense reconstruction phase. In figure 2 we can see the feature detection in images (Fig. 2A) and the overall matching between images (Fig. 2B) here represented as a
CMVS/PMVS starts by clustering images in smaller sets that can be processed within the available RAM. Then, these groups of images, called options, are processed one at a time to retrieve dense point clouds. All the point clouds are in the same coordinate frame and have the same scale, but additional ground control points are needed to scale the final model to its real size. Table 1 shows some relevant data about this case study, namely about data recorded, hardware and software used, and processing parameters.

It is important to notice that the spatial resolution of the reconstructions is also dependent of the camera distance to the objects, meaning that closer distances allow denser reconstructions but require more images to cover similar areas. The distances from camera to object ranged from 2m to 15m. A very dense point cloud model was obtained, with more than one point per square millimetre (Fig. 3), with approximately 130 million points.

Scale and orientation were done against a terrestrial laser scanning point cloud. That survey was done with a time-of-flight (TOF) scanner with an average point density of 5mm. Twenty one homologous points between the two models were selected. From those points, scale and orientation was recovered. The registration error was about 9mm. In part, this error is due to manual point selection over reflectance laser data. Then a mesh was generated from the TLS point cloud. This mesh was set as a reference to inspect the ADP point cloud. This task was performed using MeshLab software. Before inspection, the ADP point cloud was downsampled to approximately 5mm point spacing. Using a colour code, distances were mapped in the reference mesh (Fig. 4).

Visual analysis of the results shows that generally the differences are below one centimeter. Bigger differences occur in the areas where the TLS mesh resulted from interpolation of points over void areas. This shows that an image based approach can provide reliable and
accurate data when compared with TLS data assumed as ground truth.

3.2 “Pátio dos Carrascos”

The need for this survey came from a demand for stratigraphical analysis in an area where an architectural intervention is being planned. An ancient pavement was uncovered with the excavation and recording was needed.

The archaeological recording is usually done stone by stone. In order to do this, a large number of points have to be identified for proper depiction of the features of each element. This can be easily done over images, for instance rectified images. However, for three dimensional objects, this technique is not very suited. In this case, ortho images are more adequate, but they require a 3D model from which to be extracted. And this, until recently, was a very time consuming task and had to be done manually or semi-automatically. To overcome this issue, we used the ADP (SFM/MVS) approach. In table 2 it is presented some data about the processing of this case study.

A 3D point cloud model was generated with approximately eight million points over an area of 110 square meters, corresponding to an average point density of seven points per square centimetre. The generated point cloud model was then scaled and oriented using four topographic control points measured with a total station and manually identified in the point cloud model. Small ink marks were painted on site to make points identification easier for the operator. The registration error was under 9mm. After this step, an orthoimage was exported and delivered to the archaeologists that used it to draw the relevant features for their analysis as it is shown in figure 5.

Table 2. Metadata about “Pátio dos Carrascos” case study.

<table>
<thead>
<tr>
<th>Metadata: “Pátio dos Carrascos” case study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Camera</strong></td>
</tr>
<tr>
<td><strong>Number of images</strong></td>
</tr>
<tr>
<td><strong>Image size</strong></td>
</tr>
<tr>
<td><strong>Image format</strong></td>
</tr>
<tr>
<td><strong>Platform used</strong></td>
</tr>
<tr>
<td><strong>Hardware</strong></td>
</tr>
<tr>
<td><strong>Software</strong></td>
</tr>
<tr>
<td><strong>PMVS parameters</strong></td>
</tr>
<tr>
<td><strong>Processing time</strong></td>
</tr>
<tr>
<td><strong>Point cloud generated</strong></td>
</tr>
</tbody>
</table>
This example allowed us to demonstrate that the orthoimage that resulted from the automatic photogrammetric processing has both the metric and semantic qualities (as defined above) needed for the archaeological recording and analysis, that is, the archaeologists were able to recover from those images the significant data required to perform the stratigraphical analysis and the level of detail of the drawing was similar to what could be obtained by traditional techniques.

3.3 “Paços do Infante”

Also in the case of “Paços do Infante”, recording was needed because architectural intervention is under planning. A XXth century structure was dismantled uncovering XVth century walls and pavement (Dias 2012). Since the walls were approximately 10m high, a traditional survey and recording would take much time and would require a scaffold. Rectified imagery would be a possible solution but only for the walls. So it was decided to essay the SFM/MVS approach. In order to enhance image quality, reducing the strong contrast between shadow and lighted areas caused by the sunlight, bracketing was used during field acquisition of the photographs. Then, using the open source software Hugin, alignment of the images was done with align_image_stack command-line tool and exposure fusion was performed with enfuse tool (see http://enblend.sourceforge.net/enfuse_details.htm for more details). In table 3 some data on the processing of this case study is presented.

By decreasing the threshold and Wsize parameters and increasing the quad parameter, more points can be reconstructed but positional quality of the points may decrease. By decreasing the minImageNum to the value of 2, that means that it is enough that a point is only visible in two images to be reconstructed. With this the total number of points suffers a significant increase. But this also has a negative impact in the positional quality of the points. Nevertheless, sometimes this is the only way to minimize voids in the point cloud model. Changing the parameters also impacts the total time of reconstruction, as shown in Table 3.

<table>
<thead>
<tr>
<th>Metadata: “Paços do Infante” case study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera</td>
</tr>
<tr>
<td>Number of images</td>
</tr>
<tr>
<td>Image size</td>
</tr>
<tr>
<td>Image format</td>
</tr>
<tr>
<td>Platform used</td>
</tr>
<tr>
<td>Hardware</td>
</tr>
<tr>
<td>Software</td>
</tr>
<tr>
<td>PMVS parameters</td>
</tr>
<tr>
<td>Processing time</td>
</tr>
<tr>
<td>Point cloud generated</td>
</tr>
</tbody>
</table>

Table 3. Metadata about “Paços do Infante” case study.
The final point cloud model was scaled and aligned to a previously TLS survey of the site, made before excavation. Four homologous points were manually identified in both point cloud models in areas that remained unchanged during excavation. In figure 6 we have represented the site before excavation (Fig. 6A) where a TLS survey was done, after excavation (Fig. 6B) where a ADP survey was done, and the overlapping of both surveys (Figs. 6C and 6D).

The error of this registration was approximately 2cm. This may be explained by the chosen combination of parameters, particularly the “minImageNum”, leading to a more noisy reconstruction with less positional quality. Nevertheless, the overall quality of the model was suited to generate ortho images of the plans, at the scale needed, to be analysed by the archaeologists. The overlapping of both surveys also enabled a visual perception of the amount of excavation that was done, and of the position of the uncovered features (Fig. 6D).

4. Conclusions

It can be suggested that image based three dimensional reconstruction techniques (ADP), such as the ones referred, can present themselves as effective and low time consuming ways of maintaining a site recording always up to date, using only standard uncalibrated digital cameras for data acquisition. It should be added that the three dimensional reconstruction is almost unattended and automatic, what doesn’t happen with the traditional photogrammetric approaches.

One the advantages of an image based approach over terrestrial laser scanning is that it is more versatile, and in our opinion the handling of a camera is much easier even for non experts than the handling of a scanner. For instance, a camera can be easily placed in a mast, in a drone to record higher parts of a building, an underwater camera can be used in rivers or ocean sites, a macro lens can be used for smaller scale objects, and photos can be immediately processed, can be transmitted and processed at a distance, or just archived for later processing and historic reference (without the need for immediate processing). One of the disadvantages is that it is always needed extra information to scale and orient the models. This extra information can be collected with other measuring systems.

In our opinion, the possibility of recovering the geometry and texture of a site, with a method that uses standard photographic cameras, that can be easily learned, and is extremely versatile regarding the use scenarios, represents a new field of opportunities, specially for the low budget archaeological recording and analysis. Regarding the image processing, the availability of freeware, open source software and free online services, represents a democratization of the access to three dimensional processing that was not possible before.

With the appropriate setting of parameters it was shown that SFM/MVS has an acceptable metric performance when compared to TLS as it was shown with the first case study where discrepancies were generally under one centimeter. The fact that colour generation in ADP point cloud models is part of the overall geometric processing also represents an advantage when comparing with TLS. Colour is a very important quality of the data generated (ortho images for instance) for the archaeologist interpretation and analysis. Notice that in TLS colour acquisition and mapping is always an extra step of the processing.

The issue with these techniques is the hardware requirements, namely the processing performance of the computers, available RAM in the GPU (when using software that makes use of the GPU - the case of VSFM) and available RAM in the operating system (from 8GB for the smaller projects, to as many as 48GB or even 96GB for the bigger projects).
Acknowledgements

The authors wish to thank to FCT (Foundation for Science and Technology) for funding the research where the present work was integrated, namely the research project “Contributions to Architectural Heritage Conservation: Documentary Methodology based in terrestrial digital photogrammetry and 3D Laser scanning”, ref.: PTDC/AUR/66476/2006.

References


Photographic Rectification and Photogrammetric Methodology Applied to the Study of Construction Process of Provincial Forum of Tarraco

M. Serena Vinci
Institut Català d’Arqueologia Clàssica, Spain

Abstract:
The purpose of this paper refers to the application of the photographic rectification and photogrammetry methodology used as graphic support for the study of the constructive processes of Roman structures. The analysis of construction processes of the Provincial Forum of Tarraco, raises the need for a base of accurate and essential graphic documentation, which realization at the same time, meets the limitations of different types (space, time, financial resources, etc.) often impose the study of urban archaeological evidence. Depending on the location of the structure and the amount of data that it could have chosen to provide employment for a refund a two-dimensional or digital model of it. The combination of using ortho-images obtained with the software PhoToPlan and obtaining digital models of the structures by the program Topcon Image Master, have been a perfect solution to get the results of our interest.

Keywords:
Photographic Rectification, Terrestrial Photogrammetry, Roman Construction Processes, Provincial Forum of Tarraco

1. Introduction

The advantages of the use of photographic rectification and photogrammetric methodology for the documentation of the architectural heritage has been widely proved. Digital photogrammetry and photographic rectification techniques can provide straight and affordable alternative solutions for the geometric documentation. Indeed, the use of high resolution digital camera, cheap and easy to use, in combination with geometrics measurements, with or without the need for topographies controls points, permits to achieve the collection of the required data for the documentation of the monument.

As a part of a more wide-ranging research on going aiming to elucidate the technological aspect and the organization of roman construction, starting with the direct analysis of archaeological evidence, herein it is described the application of both methods used as graphic support for understanding of the construction processes of the Provincial Forum of Tarraco. The analysis of such construction processes required of an accurate and essential graphic documentation, something particular difficult to produce for a city like Tarragona: it is a city by time of continuous habitation, it developed within the same urban space since the late antiquity with the subsequent reuse of roman structure. This circumstance affected the preservation of archaeological remains and influenced the documentation of the structures included in the building fabric of modern city. In this context, depending on the limitations typical of the urban environment (space, time, financial and technical resources, etc.) and on the relevance and the amount of available data, a two-dimensional reproduction or a digital model was used.

The terms of Provincial Forum (Ruiz de Arbulo 2007) refer to huge imperial cult complex raised above the acropolis of the Roman colony since the Flavian era (Fig. 1). The monument represents an exceptional construction into the provincial capital of Hispania Citerioris; it was located in a prominent position of the colony, referred to as “parte alta”, away from the
commercial, political and religious city center of the Republican era, with a favourite scenic location and the spectacular visual impact for those arriving by sea.

The complex was divided into three terraces with axial disposition: the upper square, or “Recinto de Culto” was a themenos porch, square and surrounding the temple probably dedicated to the princeps (Pensabene-Mar 2010; Macias 2011); the intermediate terrace o “Plaza de Representación” was the seat of the Concilium, it was surrounded by a podium and two overlapping cryptoporticos. The epigraphic evidences (Alföldy 1973) dating both squares to Vespasian era. However, the lower terrace, occupied by the circus, is dated to Domitian period. The originality of this building is determined by its transversal position compared to the rest of the complex; at the same time it constitutes a sort of limit with the rest of the urban space colony.

2. Objectives

The application of modern informatics technologies for the study of Provincial Forum, aimed to provide an accurate graphic documentation for the study of construction processes and techniques. The methodology applied to this study can be divided in two main parts:

- the descriptive documentation with specific file of construction technique;
- the graphic documentation.

The use of ortho-images and digital models led to obtain a large amount of information in a fast, effective and economical way; moreover they have facilitated the interpretation of the architectonical stratigraphy and the constructive analysis of Roman structures.

The main objective of the use of photographic rectification and photogrammetry methodologies, was to achieve a compromise which is suitable for the specific context of our work.

The research project concerning the study of construction process of Provincial Forum, using two-dimensional real images, perfectly scaled or digital models, had as main purposes:

- the determination of stratigraphic architecture and construction belong to the Roman era (the first step of the research was the direct analysis of the structures; subsequently, we define the limits of the different building actions for each wall facing on the images obtained and we determine the relative chronology).
- the definition of Roman construction process (recording the construction variants identified through the analysis of building techniques; studying modulation using in the building of the structures; analysing the processing of building materials, especially of the “construction en pierre” (Bessac 1986).
- the diagnosis of structures for restoration or conservation.

3. Methodology

The software used for this work were PhoToPlan (Kubit) and Image Master (Topcon). Depending on the location of the
structure, on the interest and on the amount of available data, a two-dimensional reproduction or digital model was used. Photoplan works as extension of Autocad and it is a software for the creation of ortho-images, while Image Master is a photogrammetry software that can produces digital models or scale-plan image. In our case, both programs were used according with the different needs we met: the location of the structures combined with the adaptability and quickness for the data acquisition and data processing, were evaluated in software choice. Indeed, the study of architectural heritage in the urban context requires carefully consider current location of archaeological remains because it affects the access and the times of the intervention.

3.1 Image Acquisition

Both programs offer the possibility to choose between two methods of work: data processing with or without topographic control points. In both cases the result is a scaled image that connects photographic documentation with exact geometrical information.

Photos can be created using any type of camera. For applications proposed here was used the Nikon D-80, a digital high resolution camera, with a width of 18 mm focal length. The focal range of this magnitude was an excellent compromise between a wide field of view, to photograph the largest surface possible for each frame, and the ability of the programs used to process the images properly. In fact, the larger the field of view, the lower the amount of photos needed to document an entire surface. On the other hand, the use of a wide-angle lens produces a greater optical deformation of the photo (especially the side margins) and this makes it difficult, if not impossible, to obtain good rectification results. In our case, the focal length used was the maximum capacity that could support both programs to correctly process the data. Afterwards the take of the photos is acquired according to specific software requirements. Photoplan requires:

- one front photo for each surface by rectify,
- the position of the image do not need to be strictly perpendicular but can have a trajectory oblique always from the bottom up (this is an advantage because it avoids the use of extendable arms in the case of the documentation of large structures).

Image Master requires:

- stereoscopic pairs of photos for each surface to be rectified,
- front and perpendicular pictures.

Following the take of the photos the image is acquired by software and the camera must be calibrated with the lens used for image capture. The two programs follow a different procedure for the calibration of the camera that in both cases creates a file that contains the calibration parameters useful to being processing of the image.

3.2 Data Processing: Case Studies.

3.2.1 Photographic rectification with topographic control points: software Photoplan

Project 1. (Fig. 2) In this case the choice of the process of data acquisition and processing is mainly due to the location of the structure.

Type of structure and current location:

- west wall facing of the “sala axial” of the roman forum (rectangular hall of the north portico of the upper square),
- open courtyard unused,
- good depth of field.

Data acquisition:

- 1 photo,
- 6 topographic control points (with total

2 The program needs a minimum of 4 topographic control points,
station Topcon GPT-7000i series),

Data processing:

- loading of the image and the control points in Autocad
- connections between the points in the image and topographic points.

The good depth of field permits to photograph the structure in a single frame without compromising on quality and good resolution of the image. In this case the elements of easy identification on the structure have been chosen as control points.

For this project Photoplan revealed to be a fast and effectively method for acquiring and processing the data. Hence, according to the building location, this is the most effective and straight way to operate.

One of the results obtained from the documentation produced, was to study the traces of the work tools used to elaborate the materials of construction. For example, in the lower part of the facing are visible an oblique and parallel traces with which it is possible to follow the movement of cantero and the elaborating mode of the blocks (Fig. 3).

**Project 2.** (Fig. 4) In this case the choice is due to highest quickness of the software for the images and data acquisition.

Type of structure and current location:

- south wall facing referred as “Torre del Pretorio”\(^3\) (located in the south-east corner of the middle terrace of the forum, it connected such terrace with upper themenos),
- building currently seat of the Museum of

but it is always better to use more than 4 points, because if one of them is not exactly determined, the software will perform a statistical compensation with other recorded points. However, it is not useful determine more than 8 control points for each picture because, due to self-compensation, more points of reference mean more chances to accumulate a higher margin of error.

\(^3\) The height of the facade is about 23 m.
History and located in a town square,

- very few depth of field⁴.

Data acquisition:

- 25 Photos,
- 6 topographic control point for frame.

Data processing:

- loading of the image and the control points in Autocad,
- connection between the points in the image and topographic points for a photo-mosaic creation.

The size of the wall facing combined with a few depth of field imposes to take many photos to document the entire structure. The use of photogrammetry method would have double the number of frames to achieve (a pair of photos for each surface by rectify) and to process; in this case, considering our purposes, it was more convenient to apply the photographic rectification with control points and to process them in the program PhoToPlan.

The so-called “Torre del Pretorio” was one of two communication towers of the forum. The building has suffered numerous transformations in the past periods, and represents a significant example of the evolution of Roman architecture in the city. In addition, the southern facade presents interesting archaeological aspects because preserves part of the east wall of the middle terrace, decorated with pilasters and capitals of the Tuscan order. The study of the monument began with the interpretation of the stratigraphy architecture, after focused on the analysis of Roman construction.

3.2.2 rectification with geometric method (without topographic control points): software Photoplan

Project 3. (Fig.5) In this case the choice of the process is determined by the small size of the preserved remains.

Type of structure and current location:

- wall facing of the rectangular exedra of the east portico in the upper square of Provincial Forum,
- exhibition hall in a museum (Museu Bíblic Tarraconense),
- small depth of field⁵.

Data acquisition:

- 2 photos,
- 3 reference measurements (two measurements perpendicular to each other⁶ to allow the program to calculate the ratio between width and height of surface. The third measure is a diagonal, which must pass through most of the structure).

Data processing:

- loading the images,

---

⁴ It places at a distance between 5 and 6 m from the building front.

⁵ To approximately 1 m from the Roman wall is placed a marble reproduction of the temple of Jerusalem which does not allow a single picture of the structure.

⁶ Two orthogonal lines were determined with the use of a portable laser level, marking their ends with colored plasticine.
Photographic Rectification and Photogrammetric Methodology Applied to the Study of Construction

Maria Serena Vinci

3.2.3 Photographic rectification with geometric method (without topographic control points): software Topcon-Image Master

Project 4. (Fig. 6) This project shows a case where it is not possible to apply the methods described above.

Type of structure and current location:
• subsoil of the cathedral cloisters,
• window of the west porch of the upper square,
• small depth of field.

Data acquisition:
• 8 stereoscopic pairs of photos,
• 6 reference measurements (arranged along the edges of the photo).

Data processing:
• loading the images,
• creation of stereoscopic pairs of photos,
• union of the images in a single ortho-mosaic.

The choice of this method has been conditioned by the space in which to work. In fact the area is accessible only through a walkway that is located approximately 1.50 m from the Roman wall (Fig. 7). In this case the use of the total station is difficult due to an inappropriately support surface, i.e. the metal walkway, which is too sensitive to motion. Under the boardwalk, where the excavation area is left exposed, the depth of field increases slightly (2.50 m) but the area is occupied by large blocks of stone, found in situ, which prevent the placement of the total station. Similarly, the use of the geometric method with the software PhoToPlan is inconvenient because

reaches about 2.50 m.

8 The software screen provides a contemporary view of the pair of pictures. The program also in the picture right identifies the points identified in the picture on the left automatically, provided that they are easily recognizable.
of the difficulty of measuring the diagonal of the structure. The software chosen in this case provides a good compromise between time and flexibility for data acquisition.

The precise documentation of all the windows of the Forum allowed to compare with them about the formal and structural aspect and construction technique.

3.2.4 Digital photogrammetry: software Topcon-Image Master

Project 5. (Fig. 8) In this case the purpose is to realize a digital model of the structure that is conserved in three dimensions.

Type of structure and current location:

- north wall and part of the east wall of the “Plaza de Representación” of the forum,
- a town square,
- good depth of field.

Data acquisition:

- 12 stereoscopic pairs of photos (24 photos),
- 40 topographic control points.

Data processing:

- creation of stereoscopic pairs of photos,
- drawing of the perimeter of the facade and of the distribution of relief and voids in order to obtain a three-dimensional model.

The realization of the model is divided into 4 projects, one for each side of the structure. For each model is necessary to establish at least 3 check points (it is best to locate 4 control points in case one of the 3 has a high margin of error); the connection system between the image and topographical points is equal to that shown by the geometrical method.

In this case it is interesting result reproduce a digital model because the structure is conserved in three dimensions, even if only partially. Moreover, the use of the system with control points allows locate the structure with real topographies coordinates and connects it with the remaining parts of the forum which will be virtually reproduced.

4. Results

The project shows the advantages of the use of photogrammetry and rectified photography for documentation of the techniques and processes to build the forum Provincial Tarraco. In both cases, of course, a result at high resolution and geometric accuracy was obtained.

For the production of ortho photos, it was always chosen the faster method for the acquisition of data. In the cases of so-called “sala axial” and the “Torre del Pretorio”, the use of software PhoToPlan with control points, has allowed to obtain optimal detail-rich images.

Moreover, according to our experience, the use of the two programs proved to be an excellent solution for the particular working environment we worked in. In the cases of projects 3 and 4, the use of photoplan and Image Master without control points, was the perfect solution due to the location and size of the structure.
Finally, the construction of digital models with the photogrammetric method, allowed to obtain a reproduction of structures that are conserved in three dimensions. Is well known that as 3D support, laser scanning is a powerful technology capable of collecting vast number of surface points in far shorter time. However, apart from the high cost, a laser scanning approach faces problems of post-processing data.

The rectified images and 3d modelling also provides useful information on morphology or deformation, being a tool in its own right in evaluation and restoration processes. Indeed a precise documentation of cultural heritage status is essential for its protection and scientific studies carried out during the restoration and renovation process.

References


Image-Based 3D Documentation of Archaeological Trenches Considering Spatial, Temporal and Semantic Aspects

Robert Wulff and Reinhard Koch
Christian Albrechts University of Kiel, Germany

Abstract:
This paper presents a 3D documentation system for archaeological trenches. Structure-from-Motion (SfM) techniques are used to reconstruct the 3D surface geometry of a trench based on a set of images. Since an excavation is a dynamic process each layer needs to be reconstructed individually. A new algorithm is proposed to align the reconstructed 3D models of the individual layers automatically by exploiting knowledge about the trench geometry. This alignment enables archaeologists to reproduce the development of the excavation over time virtually in 3D space. The 3D models are classified semantically according to the archaeological entities present in the trench. To achieve this CAD plans are utilised which are already part of the standard documentation procedure in archaeology. The classification of the 3D models is finally used to semantically classify photos registered against the models as well.

Keywords:
Virtual Archaeology, 3D Documentation, Structure-From-Motion, 3D Model Alignment, Semantic Classification

1. Introduction

Documentation plays an important role in archaeology as the configuration of finds and features is usually destroyed by excavating them. To save time in the field, the interpretation is mostly performed in retrospect after the excavation. Therefore, archaeologists rely on accurate and efficient (in time and costs) techniques for documenting trenches.

The traditional 2D documentation methods, e.g. photos, rectified photos and CAD plans (see Fig. 1), are still widely used in archaeology, but 3D techniques have become more and more important in this field of research and provide a number of benefits over the conventional methods:

• There is no loss of information through a dimension reduction—the geometry of the finds and features is represented in 3D space.

• Users are not bound to the photographer’s choice of viewpoint and can freely navigate the scene to view it from any direction. This also allows creating true orthographic projections from the bird’s eye view (or any other direction). Such orthographic projections are commonly approximated by rectifying photographs using a homography (refer to Fig. 1).

• Correlating the reconstructed 3D models with other spatial data, e.g. distributions of small finds surveyed with a total station, is straightforward. Measuring distances and volumes in 3D space becomes possible as well.

• 3D models are a natural representation for the human visual system. This can be helpful when presenting results to other scientists or laymen.

Several techniques have been developed to acquire 3D models of real-world scenes, with terrestrial laser scanning and image-based Structure-from-Motion (SfM) (Hartley and Zisserman 2004; Szeliski 2010) being the most prominent ones in the field of archaeology. In this work the latter approach is favoured,
because laser scanners still have high cost implications and SfM approaches require only equipment that is already part of the standard documentation procedure in archaeology, that is, a digital camera and a computer (and optionally a total station to register the models against the site’s coordinate system). A quantitative comparison between laser scanning and SfM was carried out in (Doneus et al. 2011) with the result that both techniques have about the same accuracy. In the work of (Wulff et al. 2010) estimated depth maps were compared with ground-truth depth maps to evaluate the accuracy of SfM approaches. The results show as well that high-quality reconstructions are possible using SfM.

In this work the spatial aspects are considered by reconstructing archaeological sites in 3D using SfM techniques. This has already been proposed over ten years ago, for example in (Cosmas et al. 2001) and (Pollefeys et al. 2001). In recent time applying these techniques in archaeology for documentation purposes has become more and more popular. This development has been accelerated by improved algorithms and the release of commercial applications, such as Autodesk 123D Catch (http://123dcatch.com/) and Agisoft PhotoScan (http://www.agisoft.ru/products/photoscan/), as well as open source implementations, with the combination of Bundler (Snavely et al. 2006) and Patch-based Multi-view Stereo (Furukawa and Ponce 2010) as the most prominent example. Some recent work using these readily available software or own implementations in archaeological research are (Wulff et al. 2009; Doneus et al. 2011; Ducke et al. 2011). However, the reconstructions computed in these projects only consider the geometric information of a scene, that is, the spatial aspects. But there is more to an excavation: it is a dynamic process. The appearances of the trenches change over time as the layers are unveiled one after another. Each of the single layers has to be documented individually to allow archaeologists to virtually reproduce the excavation. Using textured 3D models provides a natural colour representation of the surface characteristics. But often the interpretation of the configuration of the finds and features is simplified by using abstract representations that reduce the information to relevant entities (as in the CAD plan depicted in Fig. 1). So in order to be useful for archaeologists as a documentation tool it is necessary to consider temporal and semantic aspects as well.

As the main contribution of this paper, the temporal aspects are introduced by proposing a method for aligning the reconstructed 3D models of individual trench layers automatically. Since the single models represent discrete time steps of the excavation process, aligning them enables archaeologists to retrace the progress over time. Though the models could be aligned or transformed into the reference coordinate system using a total station and ground control points, applying an automatic method saves time and effort during the excavation. Time-efficient documentation methods are
of high importance in archaeology, because excavations are usually carried out under narrow time constraints. The error introduced by the proposed automatic method is negligible as is shown in a quantitative evaluation. The approach exploits geometrical properties of the trenches and as far as the authors know, no methods focusing on the automatic alignment of 3D excavation layer models have been proposed before.

The semantic aspects are handled by classifying the reconstructed models into archaeological semantic entities with the result that each surface area is registered with its associated finds and features. Having a semantically classified 3D model available also allows classifying 2D photographs that are registered with the model. The importance of semantic context information has already been stressed before. (Allen et al. 2004) have suggested a 3D modelling pipeline suitable for documenting archaeological sites. However, they do not classify the model itself, but instead display additional context information such as GIS data surveyed at the site. (Manferdini and Remondino 2010) use a 3D modelling software to classify reconstructed 3D models manually. In contrast to these works, here the 3D geometry itself is classified automatically on the basis of CAD plans such as depicted in Fig. 1. These plans outline the archaeological entities with polygons and are produced on-site using a total station. Since they are already part of the standard documentation procedure, they can be reused here.

The paper is structured as follows: Section 2 elucidates the required input data and how these are acquired. The 3D reconstruction pipeline developed in this project is described in Section 3. Section 4 explains how the individual models of the different layers are aligned with each other automatically. The semantic classification is outlined in Section 5. Finally, the paper is concluded in Section 6.

2. Data Acquisition

The proposed system follows the SfM approach. Therefore, a set of images is required as input. The images need to overlap sufficiently (about 80%) and can be obtained with an off-the-shelf consumer camera. The trenches documented in this project were recorded at the excavation site of Bruszczewo in Poland—see (Czebreszuk and Müller 2004) and (Müller et al. 2010) for more details about the site—during the campaign of 2008 and had a size of approximately 3x4 meters. About 16–20 images were taken from all angles around the trench for each individual layer.

The 3D reconstruction pipeline assumes the intrinsic camera parameters (focal length, principal point, pixel aspect ratio, skew and radial lens distortion coefficients) to be known and constant. These are acquired using the camera calibration method presented in (Schiller et al. 2008). As long as the intrinsic parameters remain static, the calibration can be performed before or even after the excavation.

In order to allow measuring and correlating the resulting 3D models with other spatial data it is imperative to transform them into the reference coordinate system used at the excavation site. This is achieved by placing markers, called ground control points, in the trench that are clearly visible in the images. The positions of the markers in 3D space can be measured with high accuracy using a total station. The projected 2D image positions of the markers in the images are identified manually. As the models of the individual layers are aligned automatically (see Sect. 4) this procedure is only necessary for one layer.

To cover semantic aspects in the reconstructed 3D models, CAD plans are used. These vector drawings are produced directly on-site using a total station. They represent the semantic aspects of a trench by outlining the different archaeological entities with polygons. Since these plans are already part
of the common documentation procedure in archaeology, they can be reused here.

3. Trench Reconstruction from Images

The trench reconstruction is performed after the excavation. It follows the idea of Structure-from-Motion, that is, a textured 3D model is computed from a set of images. The models are represented by triangle meshes and for each individual layer one such 3D model is generated. Fig. 3 visualises the pipeline.

The first step in the proposed reconstruction pipeline is the preprocessing of the input images. This involves converting them to grey scale (as some of the following algorithms require these), scaling them down (an edge length of 1000 pixels is usually enough) and compensation lens distortion (the coefficients are obtained by the camera calibration procedure). To minimise aliasing effects, downsampling and lens distortion compensation are performed at once.

In each of the grey scale images SIFT keypoints (Lowe 2004) are detected. This algorithm is commonly used in SfM applications, because SIFT keypoints provide invariance against rotation and scale in the 2D image space. This enables to compensate 3D perspective transformations up to a certain degree and is important in this field of application as each layer is photographed from different viewpoints. The invariance against changes in illumination allows the images to be taken with varying aperture and exposure time and therefore eases data acquisition. In this work the SiftGPU implementation (Wu 2007) is used to perform the task of detecting keypoints efficiently on graphics hardware. For each image about 10,000 keypoints are found. This high number results in a very robust and stable reconstruction. The matching is performed with the SiftGPU library as well, so runs efficiently on the GPU. This allows matching all possible image pairs with each other in reasonable time. Matching each image with all others doesn’t require any special ordering of the input sequence and therefore provides more flexibility. For each image the 5 best matching images based on the number of keypoint correspondences are selected. Geometric consistency of the keypoint
matches of the selected pairs is assured using the principles of epipolar geometry: to get rid of outliers the essential matrix (Hartley and Zisserman 2004) is computed on these image pairs using a RanSaC (Fischler and Bolles 1981) approach. Pairs with an insufficient number of keypoint matches are discarded. As a result one obtains up to 2000 keypoint correspondences between image pairs, depending on their overlap and change in perspective.

In the next step of the pipeline a hierarchical 3D scene reconstruction algorithm (Farenzena et al. 2009) is applied to compute the camera poses (positions and orientations) and a sparse cloud of 3D points. Based on the consistent keypoint correspondences the algorithm uses the principles of epipolar geometry for the reconstruction. It divides the sequence into subsequences, processes these individually and successively merges them until the whole sequence is processed. In each step there are three possible actions to take: create a new cluster from two cameras, add one view to an already existing cluster, or merge two existing clusters into one. In contrast to the original approach of Farenzena et al. a global sparse bundle adjustment (Lourakis and Argyros 2009) is applied on each cluster after each step to optimise the reconstruction. Fig. 4 shows the result of this step. The final reprojection error of that scene is 0.16 pixels, so high-quality reconstructions can be obtained with this approach. Another benefit is that camera paths are closed implicitly and don’t require any explicit treatment as for example in (Wulff et al. 2009).

Camera poses and 3D point cloud are reconstructed in an arbitrary coordinate system. To allow measuring and correlating the models with other spatial data is is necessary to transform them into the reference coordinate system used at the excavation site. Sect. 4 describes how the individual models of the layers are aligned automatically with each other, so it is sufficient to transform only one layer model explicitly. The other models are then aligned relative to that one. To transform the model, the ground control points are used. Given their 2D position in the images and the estimated camera poses 3D points can be triangulated in the model’s coordinate system using (Hartley and Sturm 1997). This gives 3D-3D correspondences as the corresponding 3D coordinates in the site’s coordinate system have been measured with a total station and are therefore known. The similarity transformation (translation, rotation and scale) between the two coordinate systems is computed using (Horn 1987) and then applied to all poses and 3D points.

At this point, camera poses and a sparse 3D point cloud have been estimated. And for one layer these have been transformed into the reference coordinate system of the site. The next steps towards a dense representation of the scene is to produce a so-called depth map for each camera. A depth map is an image that contains each pixel’s distance to the camera centre for each single pixel (Fig. 5). Several sub-steps are necessary to solve this task: First, all image pairs are rectified using cylindrical rectification (Roy et al. 1997), as this technique is applicable to all camera configurations. Having rectified image pairs is a requirement for disparity estimation, because it reduces the dense correspondence problem to a 1D search problem. The disparities are estimated using semi-global matching (Hirschmüller 2008). The resulting disparity maps often contain frayed areas at discontinuities. These

Figure 4. Estimated camera poses and 3D points. The cameras are depicted as pyramids while the point cloud is a sparse reconstruction of the trench surface. The final reprojection error of this reconstruction is 0.16 pixels.
are removed by applying a rank filter: For each pixel the filter checks if the number of valid disparity values within a small window around the current pixel is above a threshold. If so, the pixel is accepted, otherwise it is discarded. After that, the disparity maps are converted to depth maps. For each image a set of depth maps is obtained, one for each pair that the images shares correspondences with. These are fused into a single depth map for each view by applying a median filter on the depth values of every single pixel.

All the depth maps are then combined into a volumetric representation similar to the approach of (Pulli et al. 1997). First, an axis-aligned bounding box is defined that encompasses the volume to be reconstructed. The extent of the bounding box is either estimated from the 3D point cloud or supplied by the user to focus only on relevant parts. The volume within the bounding box is subdivided into singles voxels. For increased efficiency an OcTree is used. Each voxel is then projected into the depth maps to check if it is located below or above the surface. So for each voxel it is determined how many cameras voted for it to be below the surface and how many cameras have observed it. The ratio of these two values is a plausibility of how certain it is that the voxel is below the surface. This volumetric representation is then converted into a triangle mesh using the marching cubes algorithm (Lorensen and Cline 1987).

To colourise the triangle mesh a simple approach is followed: First, for each triangle it is determined by which cameras it was observed. Second, views in which the triangle is occluded are eliminated. This is achieved by projecting it into the depth maps and comparing the distance between it and the current camera to the depth values that its projection covers in the current depth map. If the differences between these values are above a threshold the view is discarded (note that the threshold is scene-dependent as the scale for models not transformed into the site’s coordinate system is arbitrary). The mean difference is taken as the first summand for the cost function. Third, for each triangle the angle between its normal and the optical axes of all cameras is considered. This angle is the second summand of the cost function. This enables that cameras with a nearly perpendicular view on the triangle are chosen. The weighted sum of both summands forms the cost function to choose a texture for each triangle. The resulting model is depicted in Fig. 6.

Using the proposed techniques a 3D model for each single layer is reconstructed. One of these models has been transformed into the site’s reference coordinate system while the others are aligned to that one automatically as described in the next section.

4. Automatic Layer Alignment

An excavation is not a static process. The trenches change over time as the layers are unveiled one after the other. This requires each
layer to be documented individually. Using 3D documentation techniques based on Structure-from-Motion, archaeologists can obtain an individual 3D model for every single layer, each reconstructed in its own arbitrary coordinate system. In order to correlate these individual models it is necessary to align them with each other. Using the approach described in Sec. 3 to transform the reconstruction into the site’s coordinate system requires user intervention, that is, measuring ground control points with a total station on-site, identifying them in the images and entering coordinates in the user interface. Since this is cumbersome and error-prone, an automatic method is desirable.

As was stated before, the appearance of a trench changes over time. But taking a closer look at Fig. 7 one can observe that this is mainly true only for the ground plane. The profiles remain relatively static between two layers—they only grow in depth. So the approach proposed in this paper is to establish correspondences between the keypoints located in the profiles of different layers. For a sufficient subset of the keypoints the corresponding 3D points have been reconstructed which allows establishing 3D-3D correspondences between the individual coordinate systems of the layer models. These form the basis for aligning the models with a similarity transformation between the two coordinate systems. The challenge is to locate the image areas in which the profiles are. To solve this problem, data from the reconstruction step is reused. The 3D model of the layer transformed explicitly into the reference coordinate system of the site along with its associated data produced during the reconstruction stage is taken as reference, that means all remaining layer models are aligned to this one.

First, a clustering on all normals of the reconstructed 3D model is performed to determine principle orientations of triangles in the model. For this purpose the mean shift algorithm (Keinosuke and Hostetler 1975) is used. Performing the clustering in normal space is invariant to changes in translation and scale. A change in the rotation doesn’t affect the relative configuration of the normals and therefore of the retrieved cluster centres. This
means that the arbitrary coordinate systems of the models have only a relative impact on the result. The clustering yields five clusters (see Fig. 8). As expected these correspond to the principal planes located in the trench, that is, the ground plane and the four profile planes.

For the further computations only the four clusters which centres lie in one plane are considered as these are the profiles. Each of the four clusters corresponding to one of the four profiles is processed individually: For each normal in the current cluster the corresponding triangle is searched for and its centre of mass is computed by averaging its three corner points. Then a plane represented in Hesse normal form is fit into the point set using a RanSaC (Fischler and Bolles 1981) approach. The result is a set of four ideal planes representing the four profiles (Fig. 9).

Afterwards triangles with their corner points being within a certain distance to one of the planes are determined. Fig. 10 shows an example. Since each model resides in its own arbitrary coordinate system the distance threshold needs to be set individually for each layer model. Note that the triangles found by this technique are only an approximation of the profiles. Since the planes are represented in Hesse normal form they have an unbound extent. Therefore triangles located outside of the trench, but close to a plane are considered as well. But the following steps are robust enough to deal with these outliers.

The next step is to project the triangles into the views of the input images using the estimated camera poses. The triangles are rendered in white on a black background to produce mask images. So for each pixel it is now known if it belongs to a profile (coloured in white) or not (coloured in black). The classification of the detected keypoints is straight-forward: if the current keypoint is located at a position with a white pixel it is assumed to be part of a profile, if it falls to a black pixel it is discarded. Fig. 11 shows that this techniques removes most keypoints that are irrelevant to the alignment.

After this step a set of keypoints located in the trench profiles has been determined for each individual layer model. The keypoint sets of the current and the reference reconstruction are then matched as described in Sect. 3 to obtain 2D-2D correspondences between the
different layers. Since for many keypoints the corresponding 3D points were estimated during the 3D scene reconstruction one obtains 3D-3D correspondences between the two models and therefore between the two coordinate systems. Finally, a similarity transformation is computed using (Horn 1987) combined with RanSaC (Fischler and Bolles 1981) to eliminate mismatches. As a qualitative evaluation Fig. 12 depicts two aligned layers.

To quantitatively evaluate the approach, the layers 1 and 3 have been aligned with the reference layer 2 (refer to Fig. 7 for the input data of the layers). The latter one has been transformed to the reference coordinate system explicitly using the ground control points as described in Sect. 3. The accuracy is evaluated by considering the 3D coordinates of the ground control points of the aligned layers 1 and 3. For these points the true 3D coordinates in the site’s coordinate system are known, so triangulating the 2D projections of the markers seen in the images and comparing the obtained 3D points to the known coordinates allows drawing conclusions about the accuracy of the automatic alignment.

As Table 1 shows the error introduced by the automatic alignment algorithm is very small. The running times for both alignments were each about 90 seconds (including the time for loading the large amounts of data from hard disk). As a conclusion it can be

<table>
<thead>
<tr>
<th>aligned layer</th>
<th>automatic alignment</th>
<th>explicit transformation</th>
<th>error difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>abs. error</td>
<td>rel. error</td>
<td>abs. error</td>
</tr>
<tr>
<td>1 with 2</td>
<td>1.19cm</td>
<td>0.298%</td>
<td>0.91cm</td>
</tr>
<tr>
<td>3 with 2</td>
<td>1.22cm</td>
<td>0.305%</td>
<td>0.87cm</td>
</tr>
</tbody>
</table>

Table 1. Alignment errors. The table shows the mean distances between the triangulated coordinates of the ground control points and their ground-truth coordinates measured on-site with a total station. The distances were computed for the proposed alignment algorithm and the explicit transformation described in Sect. 3. The relative errors are given with respect to the maximum trench extent of 4m. The last column depicts the difference between the errors of the proposed method and the explicit transformation. One can observe that the error for the automatic alignment slightly increased in comparison to the explicit transformation, but is negligible.
stated that the proposed approach shortens the time needed for documentation significantly compared to an explicit transformation (which encompasses surveying a set of points in the field using a total station, identifying the marker positions in the images manually and entering coordinates into a user interface), while introducing only an insignificant error.

5. Semantic Classification

In the previous step the individual 3D models for each trench layer have been aligned automatically to one reference reconstruction that resides in the site’s coordinate system. Therefore all single 3D models of the layers are now in the same absolute coordinate system used for the excavation. This forms the basis for the semantic classification method proposed in (Wulff and Koch 2011) as the CAD plans representing the semantic entities located in the trench layer are produced within the same coordinate system. The method is applied on each reconstructed 3D model and its associated CAD plan individually.

As an extension to the semantic classification of the 3D model it is also possible to classify 2D images. To demonstrate this capability the original input images are used. Please note that this approach is not limited to this set of images. Every image registered to a reconstructed model (with the same techniques...}
as described in Sect. 3) can be classified in this way as long as the intrinsic camera parameters and the external camera pose are known. It is then straight-forward to project the triangle mesh of each single entity into the view to create mask images again. These define which pixels of the image belong to which archaeological entity. See Fig. 15 for an example. Sect. 6 describes possible applications for the classified data.

6. Conclusions and Outlook

This paper presented a holistic 3D documentation system for archaeological trenches considering spatial, temporal and semantic aspects. In the first stage 3D models of individual excavation layers are computed using image-based Structure-from-Motion techniques. The model of one layer is explicitly transformed into the reference coordinate system used at the excavation site and acts as a reference reconstruction for the second stage. In that stage the individual layer models are aligned with the reference reconstruction. So all models reside in the reference coordinate system which allows measuring and correlating them with other spatial data and reproducing the excavation virtually over time. In the third stage the 3D models are classified according to the archaeological entities present in the trench. This 3D classification also allows classifying 2D photographs registered against the model.

For the future it is desirable to extend the proposed alignment algorithm to circumvent the restriction that trenches need to be delimited by perpendicular surfaces (the profiles). This would allow the proposed method to be applicable to a broader variety of trenches. Besides that the authors plan to develop a data management system in which the individual representations can be stored in a unified way. This system should allow additional context and meta information generated for example by the semantic classification to be stored as well. The individual representations of a trench (especially the photographs and reconstructed 3D models) could therefore become aware of which archaeological entities they represent. This can enhance data access in a way that a user, searching for all representations in which a certain find is visible, could be guided in that request by providing only relevant data that actually contains the demanded entity.

Acknowledgements

The authors wish to thank Jutta Kneisel from the Institute of Pre- and Protohistoric at Kiel University for her assistance and for providing data.

References


Digital Photogrammetry: a Contribution to the Study of Early Middle Ages Sarcophagi Quarries of Panzoult (Indre-et-Loire, France)

Daniel Morleghem
Université de Tours, France

Abstract:
The archaeological study of early Middle Ages sarcophagi quarries presents a number of constraints (low light, readability of tool marks, accessibility and safety), so that it requires the set up of an appropriate methodology. After testing several methods and traditional surveying tools (manual drawings, naked eye observation with artificial light…), 3D modelling and especially photogrammetry have come out as a relevant solution to carry through the study of Panzoult’s quarries (Indre-et-Loire, France). This work takes place in a Ph.D on production and diffusion of stone sarcophagi in the Southern Parisian basin. The use of three free softwares (123D Catch, Meshlab and GOM Inspect Free) has validated the technique by creating full and accurate 3D models workable in the context of our study. It also improved the methodology of shooting and processing images.

Keywords: Digital Photogrammetry, Archaeology, Early Middle Ages, Sarcophagi, Quarries

1. Introduction

Panzoult’s quarries (Indre-et-Loire, France) constitute one of the largest and best preserved of early Middle Ages stone sarcophagi quarrying centres of the Southern Parisian Basin and especially of the Loire valley (Fig. 1). Discovered in 2008, it has been studied first as part of a Master (Morleghem 2010) and now a Ph.D in Archaeology at the University of Tours, within the Archéologie et Territoires laboratory (LAT - UMR 7324 CITERES).

The understanding of the exploitation of a quarrying centre (its spatial development and its chronology, the nature and the quantity of extracted blocks…) is based on the study of the morphology and organization of each of the quarries that compose it. And also the analysis of tools marks and blocks prints visible on the walls, floor and ceiling. Usually the archaeological surveys are almost exclusively made by hand with traditional techniques, in spite of taking place in complex environments often uneven and/or labyrinthine. The observation of traces of extraction is meanwhile made to the naked eye and with artificial lighting.

Different constraints exist that limit or distort the study and therefore the understanding of these quarries: the brightness that is weak or poorly controlled; the accessibility sometimes reduced and the dangerousness of the cavities; the readability of traces of extraction because of erosion, collapse, foam or saltpetre; the distance often inadequate makes it difficult to observe and record the traces of exploitation; and the coherence of the records and drawings: a quarry plan is often done at different heights of an end to the other of the cavity (because of the topography, collapse and filling…) whereas the floor plan may sometimes be very different from the ceiling.

Confront with those difficulties the question then that emerges is to know if there is a solution to overcome the scientific, technical and security constraints? In this case, is it possible and practical to implement in the framework of our Ph.D (technically and financially)?
2. Digital Photogrammetry: a Virtual Solution to Real Problems

2.1 Objectives of the Study

3D modelling and especially photogrammetry have quickly come out as a relevant solution that can overcome the various constraints at a lower cost (Pierrot Deseilligny 2011); and by so it appear to be reachable as part of our Ph.D. With light and handy equipment, it is thus possible to easily and quickly acquire a set of photographs in narrow and dangerous spaces into which we probably go only once. It also allows achieving in a short time a 3D model accessible at all times, without gap and with a sufficient accuracy.

From June to December 2011 several attempts have been made in various quarries of Panzoult and in other cavities or subterranean quarries in the same region. The point was to test the viability of the method with our limited means and in different specific context. These tests had two aims:

- Enable a complete and accurate view of the remains in three dimensions; handing this 3D model had to be easy to allow access to areas normally inaccessible to the eye or draw;

- The 3D model acquired was also intended to make plans and sections at different levels and areas of the rooms.

2.2 The Software Suite for the Photogrammetric Data Processing

For technical and financial reasons, the choice of the softwares focused on automated solutions, free or free of copyright. Usually they do not require advanced computing skills and it is a quick and easy solution to work. Three complementary softwares particularly effective in the underground context were selected: 123D Catch, Meshlab and GOM Inspect.

The first one is 123D Catch (from Autodesk). It allows us to create 3D models from digital photographs leveraging cloud computing. The basic mesh created is then refined to a better accuracy. In a second step the model is oriented along a X, Y and Z axis and a scaling is performed. Meshlab is mainly used to merge several point clouds and models for their visualization in various forms (point clouds, mesh models and textured models). Smoother than 123D Catch for large scenes, it allows a better navigation within the 3D model. The numerous plug-in enable us to observe and analyze in detail the walls of the quarry.

GOM Inspect software (in its free version) is in turn used for the basic inspection of models: create plans and sections, redraw blocks prints and tools marks and calculate...
volumes... It offers other inspection tools and also a wide range of tools to process point clouds (cloud alignment, refining meshes, models comparison...).

2.3 Test Conditions and Equipment Used

The first tests presented in this paper were mainly performed in the subterranean quarry named Barbauderie 2 that consists of four rooms of varying sizes and shapes (Fig. 2). Natural light lights up half of them especially in rooms 1 and 4.

A mid-range Olympus E-450 with a 14-42 mm lens was used. The camera was left in automatic mode and automatic flash firing. If this choice can be criticized, it was necessary due to important variations in brightness, depth of field... and lack of experience in this field. It was also to establish a baseline to see how the photos taken qualitatively met our expectations, and if necessary to define the manual settings needed to improve their quality.

3. The Photographs: Shot and Treatment

3.1 Shooting Protocol

The first step was to establish a shooting protocol adapted to our field, the equipment used and the scientific aims. From this, perhaps more than from the quality of the photo, depend the resolution and the sharpness of the future 3D model. The varied and sometimes complex shapes of the rooms involved setting up several techniques for shooting, to adapt them to the topography (because of the filling or collapses). Three main methods have been implemented (Fig. 3):

- Axial: this method is used for quadrangular spaces: a central point is determined; the shooting is done by being back to a wall and aiming to the opposite one passing through this central point.

- Tilted back and forth: this method is suitable for linear elements. The images are taken at an angle of 45° maximum to capture a large surface of the object and to obtain a sufficient recovery of a picture to another. Once the wall covered, repeating the procedure in reverse allows completing shooting for areas not visible at first or badly exposed.

- Peripheral: unlike the previous one, this method is appropriate for objects which can be walk around, as the pillars. At least eight positions (with a gap angle of 45° minimum) are necessary to properly cover the object.

For each position three images are taken: downward, to the horizontal and upward. Some projections require selectively making closed and adapted shooting. In general the methods described above can be applied to the floor and ceiling of the quarry. The accessibility of a scene makes it sometimes difficult and their strict application and therefore requires a constant adaptation to the ground.

The overlap of shots must be between 50 and 80 % and each point of the quarry must be on at least three photographs (Fig. 4). Even only lighted with the flash the images are of good quality and usable in more than 90 % of their surface in general. The pictures background should not show too much contrast at the risk
of being unusable with the shots that constitute the foreground.

3.2 Images Processing - Assembling Photos

Assembling photos with 123D Catch has allowed correcting or completing the shooting protocol: indeed, the quality of the 3D model is not so related to the number of photos taken. The more pictures we take, the greater is the risk of deformation of the model, because of possible differences in sharpness, brightness or depth of field between shots.

Creating 3D models was done in two ways that are related to the quality of shots and the software used. The first one (called “step-by-step”) consists on assembling a first batch of a few photographs with 123D Catch and adding photos gradually in small groups. As it we can control the process, find out where the possible errors are and determine their nature. The second is to create some scenes and to join them with Meshlab: in our context, it was very useful. This phased approach allows controlling the construction of the model and eliminating poor quality pictures.

4. First Results

4.1 Objective 1: Viewing and Archiving

Example 1: Room 1

The eastern part of room 1 was the first test done (Fig. 5). Its 3D model was created from fifty photographs taken in ten minutes using the axial method. If the walls are pretty complete and accurate, floor and ceiling are not so, due to a lower photographic coverage. The model can be viewed in different ways (Fig. 6 a-d): point cloud, meshed or textured 3D model. 123D Catch and Meshlab also present colourful models, allowing a more realistic look of the quarry. No editing was made during the construction of the model on 123D Catch. After thorough verification (measurements on the field and on the model), it appears that the dimensions, shapes and angles correspond to the reality of the ground. The accuracy varies around 2 cm and rarely reaches 5 cm which is more than satisfactory. There is also no significant gap in the model. Some areas, especially in the centre of the model, that appear on a large number of images, are more detailed.
Example 2: Room 2

The room 2 is larger and more complex. Almost 140 photos were taken in twenty minutes, also using the axial method. The creation of the model was done in three stages corresponding to the three main areas of the room (Fig. 7). The overall model enable us to visualize the entire cavity or almost. The centimetre accuracy allows seeing the smallest discontinuities of the walls marking the boundaries of extracted blocks (Fig. 8). The tools marks are not always clearly visible however.

4.2 Objective 2: Topographic Analysis

GOM Inspect software allows automatically extracting sections and plans from 3D models.

It is thus possible to make in a few clicks several plans of each room at the ground-level, mid-height and at the ceiling (Fig. 9). We fully realize, overlapping them, that there are differences according to altitude in the quarry; it seems obvious that a projection of blocks prints visible on the ground or ceiling from a...
plan made to half-height would make no sense: if the number of negatives is good, the shapes and dimensions are necessarily distorted as to it.

In the cases presented here (Fig. 10), the few differences observed between the manual plan and the 3D model correspond to a mismatch between the manual plan and that of the 3D model. Indeed, the cutting axis of the 3D survey is strictly horizontal, while the manual draw is made, for issues of relevance but also time and accessibility, at different levels of the walls. The same observation can be made for section (Fig. 11).

5. Conclusion and Outlook

These first attempts at modelling quarries as part of an academic and archaeological study appear quite conclusive. With light and inexpensive equipment we have created in a short time (shooting and image processing included) complete 3D models. Although imperfect, they already offer a level of detail interesting for a preliminary study (morphology, surfaces and organization of the exploitation...). So, we can already say that our two initial objectives are met.

The techniques of shooting and the quality of the photographs remain to be improved in terms of method of shooting or for the use of flash and artificial light; other programs will be tested. It should be possible then, with 3D models even more precise, to analyze in greater detail the traces of exploitation and particularly the tools marks. Ultimately, the aim is to be able to model entire more complex quarries, too dangerous to stay and where the study would be too difficult to do.

References

question. MA Diss. Université François Rabelais de Tours.

Low-Cost Photogrammetry and 3D Scanning: the Documentation of Palaeolithic Parietal Art in El Niño Cave

Alejandro García Moreno
University of Cantabria, Spain

Diego Garate
Archaeological Museum of Biscay, Spain

Abstract:
The documentation of Prehistoric parietal art is one of the most interesting applications of 3D scanning, since it allows high resolution documentation of paintings and engravings. However, in spite of the generalization of these applications, it is still an expensive technology beyond the funding possibilities of many research projects. However, alternative low-cost methods can be implemented in order to provide small projects with limited funding access to this kind of documentation. Here we present the methodology used for documenting the Palaeolithic paintings in El Niño Cave (Spain). Using a total station and ArcGIS, a 3D photogrammetric reconstruction of these paintings was made. This reconstruction provides useful documentation for Prehistoric art study and heritage management.

Keywords:
Photogrammetry, Palaeolithic Rock Art, 3D images, Documentation

1. Introduction

In recent years, the use of 3D recording systems, such as laser scanning, has been generalized, and become a common approach in Archaeology and Heritage Management. One of these uses has been the application of such methods to record parietal Palaeolithic paintings (Groenen et al. 2008), and thus obtain high resolution 3D images of painted panels, which are used in Prehistoric Art research, heritage management or even for the construction of replicas.

However, the high cost of implementing methods such as 3D laser scanning prevents its use in cases when few funds are available. That was the case of the “Prospection, documentation and evaluation of El Niño cave Palaeolithic rock wall paintings” research project, where limited funds were available, making the application of laser scanning absolutely impossible. Due to these funding limitations, we developed a low cost methodology which would allow us to obtain a three-dimensional record of paintings, but within the available funds.

2. Location and Description of the Site

El Niño Cave (Ayna, Albacete, Spain) is one of the most important sites with Palaeolithic rock art on the Iberian plateau, where only a few of such sites have been discovered to date (Alcolea González and Balbín Behrmann 2003). Moreover, there is little evidence of Upper Palaeolithic human presence on the Meseta (Cacho et al. 2010), which makes El Niño Cave a key site for understanding the Palaeolithic occupation of the Iberian plateau.

The site is located in the Segura mountain chain, which constitutes the geological transition between the Mancha plain of central Spain and the Bética cordillera. Due to its location, the site is located on a natural corridor giving access to the Iberian plateau, both from Southern Iberia as well as from the Mediterranean coast (Serna López 1997; Davidson 1986).
The cave opens in the north-western side of Barranco del Infierno (Hell’s Canyon), a tributary of the Mundo river canyon; this area is characterized by steep relief, where vertical limestone escarpments are common; the entrance of the cave is at the foot of one of these escarpments. In consequence, it is located in a very steep landscape. However, the River Mundo canyon cuts through a high plateau of gentle relief, with a hilly environment but with some limestone peaks reaching more than 1000 m.a.s.l. Thanks to this combination, a broad diversity of environments can be founded in the surroundings of the cave; this variety is reflected in the archaeological faunal remains found at the site, as well as in its Palaeolithic paintings.

3. State of the Art and the Research Project

The Palaeolithic paintings in El Niño Cave were discovered in 1970, and published for the first time by Almagro in 1971 (Almagro Gorbea 1971), but in spite of being one of the very few Palaeolithic sites outside the classic regions of Cantabria and the Mediterranean arc, little attention was paid to it; the cave’s paintings were never reviewed in depth after Almagro’s study, and in consequence later regional revisions were based mainly on Almagro’s work. On the other hand, fieldwork directed by I. Davidson was carried out at the cave in 1973, in order to determine the existence of an archaeological deposit and its chronology; this fieldwork brought to light a sequence including evidence of human occupation of the cave during the Middle and Upper Palaeolithic as well as in the Neolithic (Higgs et al. 1976).

Because of the lack of high resolution documentation as well as a modern analysis of the paintings, we decide to revisit the site in 2008. This research was funded by the Instituto de Estudios Albacetenses “Don Juan Manuel” belonging to the Regional Government of Albacete. The main goals of this project were to review the cave walls thoroughly, searching for new paintings and/or engravings, to obtain high-resolution, digital pictures of the paintings, to create new copies of the paintings, which include a representation of wall volumes, and to interpret the paintings according to the new theoretical framework in Prehistoric art. However, limited funds prevented high-resolution 3D laser scanning, which requires large investments. Due to this limitation, photogrammetry was employed, with the intention of created a 3D image of the main panel.

4. Palaeolithic Parietal Paintings

At El Niño, prehistoric paintings are distributed throughout the cave, including an external panel of Levantine paintings in the rock shelter outside the cave, dating from the Neolithic. In contrast, all the Palaeolithic paintings are located inside the cave. There are two main sectors with Palaeolithic paintings; the first one, which includes the main panel, is located in the cave vestibule, while the second one is located inside the cave, in a small lateral gallery. Some other representations can be found in other areas of the cave, although most of the paintings are concentrated in these two panels (Garate Maidagan and García Moreno 2011).

The main panel is located in the cave vestibule, near the entrance, and it is visible in daylight. The paintings were produced on a wall on the left of the vestibule, where three different volumes can be found. The dimensions of this panel are about 250 cm. maximum length and about 105 cm. maximum height, while the centre of the panel is located about 150 cm. above the vestibule floor.

The main panel is composed of 13 graphic units, 9 of them representing figurative animals, while the other 4 represent non-figurative traces, such as lines and stains (Fig. 1). Among the figurative representations, all of them
painted in different tones of red, there are two ibex, three hinds, one horse, one indeterminate bovid and two red deer, which are the principal figures in the panel according to their central location and size. The figures show a clear spatial organization of the panel, with the two large red deer at the centre, while the two ibex are located on the sides, looking towards the exterior. Regarding the central figures, the two deer and the largest hind face the entrance (to the left), while the horse faces the inner part of the cave (to the right). Finally, the two small hinds are represented face-to-face, and in consequence one looks towards the entrance and the other towards the inner cave.

The second panel is located in a lateral gallery in the inner part of the cave, separated from the vestibule by a massive formation of speleothems. The panel is on a large, plain, inclined section of wall, although this part of the cave is in complete darkness. The panel consists of 7 graphic units, with three figurative representations and some lines and signs. The figurative representations are a small horse, a schematic ibex (only the head with large horns and the back have been represented) and a large representation of a snake, made by two parallel lines, with an indication of the head and some interior rings. If we consider the two representations of snake (cut by a calcareous concretion) as part of the same figure, it measures up to two metres long. This representation is quite exceptional, since few examples of ophidians are known in Palaeolithic rock art.

The characteristics of these paintings relate them with other neighbouring regions, mainly Southern Iberia and the Mediterranean arc, although a relationship with the Cantabrian tradition has been proposed because of certain details, such as the three-line composition of the hind heads as well as the depiction of ears as a V (Almagro Gorbea 1971), or the presence of red deer (Balbín Berhmann and Alcolea González 1994). However, these stylistic conventions can also be found in other sites in the Mediterranean arc and Southern Iberia, and in consequence the closest parallels for El Niño can be found in these areas. On the other hand, the varied composition of the representations, with such a large variety of animals, is quite exceptional, since most of the artistic ensembles are dominated by one or two animals (Garate Maidagan And García Moreno 2011).

Regarding the chronology of the paintings, a bone sample from a small archaeological level lying below the main panel produced an AMS date of 22,780±60 BP (UGAMS-7738, 27,977 – 26,934 cal BP2), which could indicate a Gravettian chronology for the paintings. This attribution is consistent with the stylistic characteristics of the paintings, as well as with other similar sites in Southern Iberia (Garate Maidagan And García Moreno 2011).

5. **Creation of a Triangulated Irregular Network (TIN)**

In the first step for obtaining a 3D image of the main rock art panel, we decided to create a digital representation of the panel surface, which would provide the support where the orthoimage could be fitted. Thanks to this support, the orthoimage will give the illusion of a 3D image, allowing the recognition of the different volumes of the rock wall.

---

---

2 OxCal 4.1, Curva IntCal’09, 20.
Because of the impossibility of using a high-resolution laser scanner due to limited funds, a total station (Leica TPS 407) equipped with Laser Distance Measure was used to record up to 1703 three-dimensional points (X, Y and Z coordinates with millimetric precision), which shaped a mass of points representing the two walls where the main panel is located (1295 points for the main panel and 408 points for a lateral, perpendicular wall where the indeterminate bovid is located).

Once these points were recorded, the resulting mass of points was used to create a continuous surface. This surface was created by interpolating the position of every single point with its neighbours, in order to generate a Triangulated Irregular Network (TIN), where each point served as node of the TIN. A TIN is formed by a series of irregular triangles, which are built by adjusting a plane to three close, non-collinear points, whose position is known; then the triangles are placed against the surface, creating a mosaic (Felicísimo Pérez 1994). By using a TIN, the heterogeneous surface of the rock wall could be modelled, and a 3D, referenced representation of the paintings’ support was obtained (Fig. 2).

6. Orthorectification

Once the TIN representing the wall where the main panel is located was obtained, an orthoimage of the panel was created. The aim of creating an orthoimage was to obtain a referenced image of the paintings with the same coordinates system used to create the TIN, in order to superimpose both layers. Moreover, the orthoimage allows measurements of figure size and position to be taken, improving the analysis of Palaeolithic paintings.

To create the orthoimage, 97 reference points were recorded using a total station. Some of these reference points were artificial ones, whose position was indicated using a removable sticker, while for some others specific areas of the paintings were selected, mainly points and line intersections (Fig. 3).

Using ArcView 3.2’s Warp Environment tool, a high-resolution picture (obtained using a Nikon D90 camera) of the main panel was georeferenced, by inserting the real three-dimensional coordinates, recorded using the total station, into their position within that picture. The orthorectification of the image was done using the Nearest Neighbour resampling method. The resulting image was a referenced picture of the main panel, resampled according to real-world rock wall coordinates (Linder 2003).

Finally, the georeferenced image was fitted to the TIN, which was possible thanks to the use of the same reference system in both layers. This way, a 3D image of the main art panel was obtained (Figs 4 and 5), which allows the image...
to be browsed and the Palaeolithic paintings to be visualised from any angle and perspective.

7. Results and Conclusions

Despite the limited funds of the El Niño Palaeolithic rock art research project, which prevented the use of a laser scanner, a 3D orthoimage of the main painted panel was obtained, using alternative instruments and procedures, such as a total station and a digital photo camera, which are available from many research institutions or academic departments. Although the resulting products were not so accurate as those obtained when using laser scanning, thanks to the use of this alternative method a three-dimensional, orthorectified image was obtained, which allows El Niño Palaeolithic rock art to be visualised from any perspective, improving the documentation, analysis and dissemination of this exceptional example of Palaeolithic rock art.

Acknowledgements

This work was made possible thanks to the funds provided by the Instituto de Estudios Albacetenses “Don Juan Manuel” of the Regional Government of Albacete, as well as the support of the Cantabria International Institute for Prehistoric Research (University of Cantabria), the Centre de Recherche et d’Etudes de l’art Préhistorique - Emile Cartailhac (University of Toulouse) and the Town of Ayna. We wish to thank Jesus Moreno, from the Ethnographic Museum and Tourist Office of Ayna, for his support.

References


3D Documentation in Archaeology: Recording Las Cuevas Site, Chiquibul Reserve, Belize

Fabrizio Galeazzi, Holley Moyes and Mark Aldenderfer
University of California at Merced, USA

Abstract:
This research aims to define new methodologies for the 3D documentation and preservation of archaeological sites. In this paper we will show one approach to document completely aspects of an archaeological site using different 3D survey technologies and find the most appropriate methods, based on diverse environmental conditions and light exposures, and with varied surfaces. During the summer 2011 fieldwork campaign a test was conducted in the Las Cuevas’s site (Chiquibul Reserve, western Belize). Thanks to this test was possible to demonstrate the reliability of the triangulation laser scanner technique in terms of accuracy in cave environment. This kind of technology allowed the high resolution data capture of the excavation process in 3D. The final result was the 3D models of the units’ levels (meshes and textures applied). Moreover the comparison between triangulation laser scanner and dense stereo matching techniques showed pros and cons of the two recording methods.

Keywords:
Digital Archaeology, 3D Documentation, Dense Stereo Reconstruction, Triangulation Laser Scanner, Digital 3D Models

1. Introduction

The research described in this paper aims to define new methodologies for the 3D documentation and preservation of archaeological sites. 3D archaeological surveys are becoming more common in archaeology, but this can become problematic because researchers have yet to integrate these technologies to develop a complete and coherent methodology for 3D documentation of sites. The proposed work is intended to completely document aspects of an archaeological site using different 3D survey technologies to find the most appropriate methods based on diverse environmental conditions and light exposures, and with varied surfaces. The final product will be the creation of a 3D application that will be used for both scientific research and the creation of models and digital objects for heritage preservation and outreach activities. In the month of June 2011 the 3D documentation methodology was tested in the Las Cuevas site, located at the Las Cuevas Research Station in the Chiquibul Reserve in western Belize (Fig. 1).

This site, originally referred as “Awe Caves”, is a medium-sized Maya ceremonial center located approximately 14 km east of Caracol. It is of particular interest because a large cave with an extensive dark zone tunnel system resides directly beneath the largest temple in the site core. This archaeological site is a perfect case study to test the different 3D documentation techniques and to integrate them in a precise working plan. The most interesting aspect of this site, for the 3D documentation, is the heterogeneity in its parts. It consists, in fact, of a number of buildings including temples, range structures, a ballcourt, and what appear to be sacbes and causeways. These characteristics represent a perfect test for the understanding of which 3D survey technologies are more appropriate for each structure category and how they can be integrated. Because of the complexity of the site, it has a wide range of environmental conditions – dark recesses of caves, areas in shaded sunlight under the jungle canopy, and areas of more direct sunlight in areas that have been cleared of brush or exposed by treefall. Thus, we have structures, variability in lighting, and other kinds of features in close proximity.
Two areas of the cave were dug and surveyed using two different approaches, the triangulation laser scanner (Minolta Vivid 910) and the photogrammetry technique (dense stereo matching). Two units have been excavated inside the cave, one in the entrance chamber, another in chamber one. The entire excavation process was acquired (9 strata in the entrance chamber and 8 strata in chamber one).

The goal of this test is the understanding if it is possible to plan the 3D documentation of an archaeological site using the cheaper and more portable photogrammetry, instead of the more expensive laser scanner technology. Can photogrammetry technique give the same result of the laser scanner in term of level of detail, or is the gap that exists between these technologies not filled yet? Aligning all the strata of the unit, it was possible to obtain the complete documentation of the excavation process in 3D. This project aims to demonstrate how these technologies can be a powerful tool for the site survey and data analysis. They give the possibility to preserve the information digitally through time. In this way archaeology can be revisited over the long-term and, thanks to the following of new discoveries, analyzed by multiple experts and subjected to new analytical techniques.

2. Related Work

In the past ten years the use of new technologies for the 3D documentation and reconstruction of cultural heritage has changed how we approach archaeological research. Archaeology, becoming even more ‘digital’, is officially part of the digital village described by Zubrow (2006: 12). The use of 3D laser scanners and photogrammetric methods is now well established in the field. One of the main reasons for this development is the possibilities that these techniques give to preserve digitally information through time.

Up to now the efforts in the 3D laser scanner documentation have been mainly focused on two aspects: the first is the 3D analysis of archaeological sites and monuments after the excavation process. Here it is possible to obtain only the static representation of a site or monument. An example is the 3D survey of Livia’s Villa (30-25 BC, Rome, Italy). A laser scanner (CyraX 2005) was used for the 3D acquisition of the site. Thanks to this technology it was possible to obtain a detailed 3D data capture of point clouds, and, after the post-processing, a mesh characterized by a millimetric precision of the entire archaeological site (Dell’Unto et al. 2008: 121-122). The main goal of this project was not the 3D documentation of the excavation process, but rather the representation of the site after the dig.

The 3D documentation of the archaeological excavation process in real-time is one of the most challenging aspects of this research. What, in fact, this work is intended to understand is if it is possible to develop a
methodology that allows the three dimensional survey of archaeological stratigraphy during the excavation process. For the achievement of such ambitious a goal, different technologies must be tested in order to integrate and use them according to their characteristics and potential.

2.1 Time of Flight Laser Scanner vs Triangulation Laser Scanner

In the past ten years the time of flight laser scanner technology has demonstrated to be very powerful in the 3D documentation of general archaeological contexts, but cannot produce sub-centimeter precision, which is often required for high quality site documentation (Galeazzi et al. 2010: 102; Koch and Kaehler 2009: 1).

For this reason in the summer of 2010 a triangulation light laser scanner, the Konica Minolta VIVID 910, was used to scan the stratigraphic units of Building 86, a mud-brick house, in Çatalhöyük, Turkey.

During the fieldwork was possible to scan 27 stratigraphic layers (Fig. 2). This first test represents a very good starting point in the analysis of the effectiveness of this new methodology.

The stratigraphy was acquired using two different laser scanners: the Konica Minolta Vivid 910 and the Trimble CX. The first one is a triangulation light laser scanner able to reach a level of detail within to a millimeter (TELE X: ± 0.22 mm, Y: ± 0.16 mm, Z: ± 0.10 mm), with a scan range of 0.6 to 2.5 m. The second is a time of flight laser scanner that can reach a detail between 8 mm and 1 cm (considering the post-processing phase), with a scan range of 0.5-350 m. Both the scanners have pros and cons.

The data recording obtained through the Minolta is not as fast as that obtained by the Trimble, which can acquire large areas in a few minutes of work. However, the data post-processing is faster; in fact the Minolta is able to acquire textures (under good light conditions) and surfaces, while the Trimble is not. The latter, in fact, can just acquire point clouds that have to be triangulated and texturized in the post-processing phase. Unfortunately, the Konica Minolta Vivid 910 cannot work in direct light conditions2. For this reason we were forced to cover the excavation area in order to scan the layers. Since the team was not prepared for this kind of situation it was not possible to have a perfect distribution of the light with the cover. So the textures acquired by the scanner were not homogeneous enough to be used in the 3D models of the layers. The textures of the layers were acquired through a high resolution digital camera in the attempt to georeference them to the 3D surfaces of the layers during the post-processing. There are two negative aspects in this kind of texturing procedure: the first one is that the alignment of the map to the 3D model is made manually through control points, and for this reason the accuracy of their matching is not always guaranteed. The second is that it is more time consuming.

The fieldwork proved that using Minolta it is possible to obtain very detailed 3D meshes in the acquisition of stratigraphic layers, and confirmed that the time of flight laser scanner cannot produce sub-centimeter precision, which is often required for high quality site documentation (Koch and Kaehler 2009: 1).

Figure 2. The 3D stratigraphic units of the settlement’s houses, B. 86, Çatalhöyük, Turkey (Konica Minolta Vivid 910 Laser Scanner).
2.2 From Photogrammetry to Dense Stereo Reconstruction Tools

The use of photogrammetry techniques for the 3D documentation of archaeological sites was tested in different digital archaeology projects. One of the most successful is the 3D documentation and reconstruction of the medieval monastery of “Santa Maria di Tergu” (X-XII century AC, Sardinia, Italy). In this project some excavation areas were 3D recorded using the photo-modelling technique (Galeazzi et al. 2007).

These techniques are used to create simplified 3D metric models with a photorealistic aspect. Photo-modelling is a photogrammetric technique that provides image-based modelling. Using it, it is possible to calculate measurements and build 3D models through digital pictures, and is used for the creation of simplified 3D metric models with photorealistic aspect.

The comparison between laser scanner and photo-modelling techniques were developed in the 3D survey of the Livia’s Villa archaeological site. In this case the high irregularity of the structures did not allow the acquisition of a good 3D resolution model through photo-modelling technique. The difference in term of level of detail between the laser scanner and the photo-modelling 3D models was considerable (Dell’Unto et al. 2006).

The use of dense stereo reconstruction tools for the 3D documentation of the archaeological stratigraphy is not anymore a new technique in archaeology. Thanks to this technique the camera calibration that was mandatory with photogrammetric software like PhotoModeler, is not necessary anymore. Dense stereo reconstruction tools, in fact, allow the 3D data generation starting from a series of un-calibrated images. The different steps of the process of 3D reconstruction are image matching, camera parameters estimation, dense matching, and the results of this computation may be similar to a series of range map, associated to each input image (Callieri et al. 2011). The processing of image sets usually takes some hours.

The main problems in the use of this technique are the lack of scale information and the absence of a convincing comparison between this documentation method and 3D scanning. Nonetheless some attempts have been made in this direction (Doneus et al. 2011; Guerra 2011), the definition of a coherent methodology is still far to come.

3. Case Study: Las Cuevas, Chiquibul Reserve, Belize

3.1 The Site

The Las Cuevas Archaeological Reconnaissance (LCAR) began investigations of the ancient Maya archaeological site of Las Cuevas, located in the Chiquibul Reserve in western Belize, Central America (Fig. 1) during a four-week summer field season in 2011. Originally referred to as “Awe Caves,” Las Cuevas has received little investigation, with the exception of one notable project. In 1957, working for the British Museum, Adrian Digby (1958) and then Commissioner of the Belize Department of Archaeology A. H. Anderson conducted excavations at the site and produced a sketch map. Digby wrote a brief article for the London News describing the site and reporting his excavations, and Anderson mentions a 1938 visit to Las Cuevas in his 1962 paper for the Americanists’ Congress, but no other reports have been found. The current investigations by the LCAR address cultural dynamics during the Late to Terminal Classic periods, a time of severe stress in the ancient Maya world, by examining the ritual life of a Maya community on the eve of collapse.

Recent studies (Aimers 2007; Demarest et al. 2004:546) question the usage of terms such as “collapse” or “fall” because they are
colorful but misleading, but Demarest and his colleagues suggest that the events of the late 9th century do represent changing political systems and ideologies. In other words, instead of representing the total failure of an entire civilization, the “collapse” has been redefined as the decline of the elite class and the abandonment of the institution of kingship in the Maya Lowlands. Most Mayanists agree that there was in fact a major change in both population and social organization, and it is clear that many sites were abandoned during the mid to late 9th century.

Las Cuevas offers an excellent venue for exploring this issue. It is a medium-sized minor administrative/ceremonial center whose nearest neighbor is the mammoth site of Caracol, located approximately 14km to the east as the crow flies. On the surface Cuevas seems to be typical of many Late Classic sites found in Belize such as Baking Pot, Floral Park, Blackman Eddy or Minanha (Iannone 2004). However, this site has something that these other sites do not—a large cave system that runs directly beneath the main plaza. The cave entrance is situated below the eastern pyramid or “shrine,” that as noted at both Caracol and Tikal, are the foci of ancestral burials (Chase 1994:53, Becker 2003:258-262). The cave entrance itself is massive, cathedral-like, and architecturally modified, suggesting that it was used for large public performances. The LCAR aims to articulate constructions in the cave with those of the surface site to produce a picture of the ritual life of the community, and contextualize ritual practice within the sociopolitical as well as the natural environments.

A large cave with an extensive dark zone tunnel system resides directly beneath Structure 1, and runs directly beneath Plaza A. The opening of the cave is at the base of the cenote located in the center of the site. While it is not unusual for Maya sites to be associated with caves, we rarely see such a direct connection or such an extensive tunnel system beneath a site core (Moyes and Brady 2012). Not only this, but located directly inside the cave’s cathedral-like entrance is an additional cenote with a natural spring at its base (Fig. 3). Test excavations both in surface contexts and within the cave were conducted to begin to establish the site’s chronology. The plan is to investigate connectivity between Las Cuevas and Caracol by comparing architectural layouts, ceramic assemblages, chronology, ritual practices and settlement patterning between the sites (Moyes et al. 2011).

3.2 Mapping the Site

The first field season of the LCAR focused on mapping the site and conducting test excavations for chronology building. The site was surveyed by Rafael Guerra, Erin Ray, and
Mark Kile, using a Sokkia 650X 6” reflectorless total station on loan from the University of California, Merced and a Topcon 3” total station on loan from Lisa Lucero. Data were displayed and organized using a Geographic Information System. Maps were produced by Justine Issavi, Lauren Phillips, Rafael Guerra, and Holley Moyes. A digital elevation model (DEM) of the site, a plan view map of the constructions in the site core and plazuela group (Fig. 4), and a partial map of the cave were created (Fig. 3; Moyes et al. 2011).

4. 3D Scanning vs Dense Stereo Reconstruction Tools

Concerning the stratigraphy data acquisition, the tests were conducted in four different areas of the site, characterized by diverse environmental conditions and light exposures, and with varied surfaces:

**Test 1 – Caves Chamber 1 (no natural light/compact and muddy soil).** The test in the second chamber of the cave allowed testing the data acquisition techniques (dense stereo matching or triangulation laser scanner) in an area characterized by the total absence of natural light. Moreover, testing this part of the cave was of extreme importance to put in evidence the performances of the different documentation technologies in high level of humidity and compact and muddy soil’s condition.

**Test 2 – Caves Entrance Chamber (medium natural light/compact and medium wet soil).** The first chamber is at the entrance of the cave for this reason it presents a medium exposition to natural light. The soil is less compact and muddy respect to the second chamber.

**Test 3 – Ballcourt (areas in shaded sunlight under the jungle canopy/wet soil).** The test in this part of the site showed the limits and potentialities of the different methods in no direct natural light conditions. The jungle canopy, in fact, permits a homogeneous distribution of the sunlight all day long.

Test 4 - Open area of the research station (direct sunlight in areas that have been cleared of brush or exposed by treefall). The test in this area was, probably, the most challenging because exposed to direct sunlight.

The tests conducted in the ballcourt (3) and in the open area of the research station (4), showed the limits of the triangulation laser scanner technique in these environmental conditions. The test 4 confirmed the results obtained in the Çatalhöyük project, the triangulation laser scanner (Konica Minolta Vivid 910) cannot work in a direct light condition. The test 3 gave the same results. In fact, also if the canopy partially filters the direct sunlight, it is still very difficult to obtain good results in these light conditions. The acquisition in these kinds of environments (3 and 4) is possible just covering the area that has to be scanned. But this procedure is not always possible during the archaeological fieldwork; moreover the textures acquired by the scanner are not homogeneous enough to be used in the 3D models of the layers. Since the triangulation laser scanner technique showed its limits in these two areas of the sites, in the summer 2012 fieldwork campaign the survey will be conducted comparing the photo dense tools techniques to a phase shift variation laser scanner, the FARO Laser Scanner Focus3D\(^3\), to understand if these technologies may work better in these environmental and lighting conditions.

4.1 Cave Investigation

The triangulation laser scanner techniques showed its limits in the outdoor environment of the site (tests 3 and 4). The result was totally different in cave environment (tests 1 and 2). The FARO Laser Scanner Focus3D uses Phase Shift technology to measure the distance to a surface. An infrared laser is sent out and reflected back to the system. The distance is measured by analyzing the shift in the wavelength of the return beam: http://www.faro.com/site/resources/download?ReturnUrl=/site/resources/download/1772/.
Two test units were placed in the cave. Cave investigations were supervised by Barbara Voorhies. Laura Kosakowsky analyzed the ceramics for chronology using standard type: variety designations largely in line with the Belize Valley (Gifford 1976). All of the units datable material and all contained ceramics dating to the Late Classic Spanish Lookout/Tepeu II complex (Moyes et al. 2011).

Unit 1 (Test 2 - Cave Entrance Chamber, medium natural light/compact and medium wet soil)

Unit 1 was of particular interest. This unit was placed in the Entrance Chamber into a partially eroded platform with a plaster floor. A second floor was encountered below suggesting that there was more than one phase of construction within the cave. Initially we thought that this earlier construction may have been quite old, but ceramic analysis clearly demonstrated that this was not the case and that the cave was modified on more than one occasion in the Late Classic period. A total of 316 sherds were excavated within the unit, of which 62 were identifiable to type. Although there were redeposited sherds from the Late Preclassic (Sierra Red Group) and Early Classic period Petén Glosswares, both constructions primarily contained sherds dating to the Spanish Lookout/Tepeu II complex. Additional artifacts encountered including chert flakes, a chert biface, and animal bone, bolster our argument that, rather than representing a unique cave assemblage, the artifacts in the fill of the platform are typical of mixed fills from surface site excavations elsewhere (Moyes et al. 2011: 17-19).

The excavation process of this unit was completely recorded in 3D using the two techniques: triangulation laser scanner, Minolta Vivid 910; and dense stereo matching. The quality of the data coming from the Minolta acquisition was very good. The characteristics of the cave environment (medium and no natural light conditions) allowed a better control on the lighting of the excavation area. Artificial lights, in fact, were used to give a more homogenous distribution of the light on the area to be scanned. The first positive aspect in the use of the triangulation laser scanner technique was the extreme detail of the meshes acquired (Fig. 6). The second benefit consists in the possibility to acquire meshes and not point clouds. This allowed, during the post-processing phase, to avoid the point clouds filtering, alignment and triangulation, saving almost the half of the post-processing work time. One of the negative aspects in the use of this technique is the low resolution of the textures. The camera integrated in the scanner, in fact, is a low resolution camera (number of output pixels: 307,000/FINE mode, 76,800/FAST mode). Moreover this technique is not recommended for scanning big areas for two
main reasons: first of all because the post-processing and alignment of the different meshes obtained from the scans will be extremely time-consuming. The Minolta, in fact, allows acquiring just a small area at a time. The optimal 3D measurement range (0.6-1.2 m) consents to record a surface of about 80x80 cm. Secondly the fieldwork experience showed that the cave environment, because the high level of humidity, remarkably affects the laser scanner performances. After about an hour and an half the hardware stopped working properly. Since in this amount of time it is possible to estimate an acquisition of a surface of about six square meters, the acquisition of a bigger area would slow down too much the excavation process.

The same unit was acquired using dense stereo reconstruction tools. The purpose of this paper is not to give an overview and comparison of the different dense stereo matching software. Some evaluations have already been done between three tools: Arc3D webservice, Photosynth/Bundler+PMVS2 and AutoDesk PhotoFly. The comparison showed that it is possible to obtain the same numerical results from one system to another. The differences between them are in terms of data density, resilience to non-optima photo dataset, visual quality of data and tools flexibility (Callieri et al. 2011). Photosynth/Bundler+PMVS2 seemed to be the best choice between the three different tools tested, because the only one that can be executed on a local machine. This represents a fundamental characteristic in archaeology, because during the on-site 3D documentation the probability to be connected to a webserver is not so common, especially in very remote site like Las Cuevas.

A different dense stereo matching software was used for the 3D documentation of the Las Cuevas stratigraphy, PhotoScan, Agisoft. This software uses the same algorithm adopted by Photosynth/Bundler+PMVS2, but it was preferred because it is the only dense stereo matching software that allows complete 3D model restitution (alignment, creation of the geometry and texture). Moreover, thanks to the graphic interface, it is possible to separately manage the geometry and texture creation from the alignment. In the last years, because of logistics issues connected to the fieldwork in remote environments, archaeologists have started to test this technique as a possible alternative to laser scanner technology. Also the site subject of this study, Las Cuevas, is located in a very remote area of the Chiquibul Reserve in western Belize. The site is four hour driving from the closest town, San Ignacio. In this kind of environment it is very difficult to bring heavy equipment like laser scanners, therefore the possibility to acquire 3D models just taking picture makes the dense stereo reconstruction tools technique extremely more flexible.

Another positive aspect in the use of this technique consists in the possibility to reduce consistently the acquisition and post-processing time. The acquisition time with the Minolta laser scanner of a surface of 2x2 meters was about twenty minutes, while the pictures capture for the dense stereo matching took about five minutes. The post-processing of the same surface with the Minolta took about an hours of work (meshes optimization and alignment), while with PhotoScan the data processing was about four hours, but the real work time, the data loading, took just about 15 minutes, the rest was of machine processing.
The final 3D models of the unit, result of the acquisition with the two techniques, have been aligned to permit their comparison. Nonetheless the Dense Stereo Tools technique gave good results in term of meshes’ details, the difference between this technique and the triangulation laser scanner technology was still relevant. The comparison of the 3D models obtained from the two techniques, in fact, showed that just the triangulation laser scanner technology allowed the preservation of the details of all the features contained in the strata (Fig. 7).

The preservation of the unit textures’ quality was problematic for both the 3D recording methods. The use of external lights remarkably affected the texture color in the Minolta data acquisition. Lighting issues, pretty common in cave environments, were evident also in the integrated camera of the laser scanner (Fig. 8). The camera flash was used instead of external lights for the dense stereo matching data acquisition in the attempt to avoid the mentioned effects on the textures’ color, but the result was exactly the same, the textures’ color of the unit was altered (Fig. 9).

**Unit 2 (Test 1 – Caves Chamber 1, no natural light/compact and muddy soil)**

This unit is located in an alcove that has a large imposing stalagmite positioned in front of the narrow passageway that leads into the alcove from the direction of the cave’s entrance. Unit 2 (1 x 1m) was located on the east side of a protruding rock that was surrounded by abundant flat-lying sherds. The diagnostics ceramics belong to the Spanish Lookout Ceramic Complex pertaining to the Late Classic Period (A.D. 700-900; Moyes et al. 2011: 19-21).

The two techniques (dense stereo tools and triangulation laser scanner) were tested also on this unit (Fig. 10), and the comparison gave exactly the same results obtained in the 3D documentation of unit 1.
5. Conclusions

The interest in archaeological 3D documentation has greatly accelerated over the past decade (e.g. Addison 2008; Bobowski et al. 2008; Dell’Unto et al. 2008; Galeazzi et al. 2007; Koch and Kaehler 2009; Zubrow 2006). Nonetheless today 3D technologies are being used more commonly in archaeology, and the use of technologies is well established for the documentation of archaeological sites (e.g. Callieri et al. 2011; Dell’Unto et al. 2008; Craig et al. 2006; Neubauer et al. 2005; Galeazzi et al. 2007), there are only a handful of scholars who compared different techniques on site, and usually this evaluations take in consideration just two technologies per time (Dell’Unto et al. 2006; Doneus et al. 2011; Koch and Kaehler 2009). This research compared the two main types of 3D laser scanner technologies (the triangulation and the time of flight) with dense stereo matching technique. The analysis of all those techniques on site is fundamental to have a complete and comprehensive test of their potentialities, and to verify the possibilities to integrate them effectively in the 3D documentation process.

The results obtained in the first Las Cuevas 3D documentation campaign showed the extreme flexibility of the dense stereo reconstruction tools technique for logistics, data acquisition and post-processing time. This technique allowed saving 15 minutes for the data acquisition (5 minutes vs 20 minutes), and 45 minutes for the data processing (15 minutes vs 60 minutes), without considering the machine processing in this estimation (about three hours and a half), respect to the triangulation laser scanner technology. For this reason it is possible to state that in this particular environment the acquisition through triangulation laser scanner technology slows down the excavation process more than the dense stereo reconstruction data recording, especially for area bigger than six square meters.

The improvement of dense stereo matching technique in term of detail, in the last years, is evident; it allowed a good reliability in the metric representation of the unit information. But it does not consent the preservation of the details of all the features in the unit. The triangulation laser scanner technique seemed more indicated for the 3D reproduction of the unit’s micro-stratigraphy in cave environment. However more tests will be done in the second Las Cuevas fieldwork campaign, in summer 2012, to understand if the camera flash light, used during the dense stereo matching acquisition, has affected the detail and quality of the 3D model’s meshes. Since the triangulation and time of flight laser scanners showed some limits for the 3D documentation of archaeological stratigraphy, the first for the impossibility to work in direct light conditions and the second for lack of details of the strata’s features acquired in the 3D models, in summer 2012 the dense stereo reconstruction tools technique will be compared with a phase shift variation laser scanner to understand if this technology could serve better the scope.

This project is a work in progress; the next fieldwork campaigns will allow its completion. But the preliminary results coming from the summer 2011 fieldwork season clearly showed the potentiality of this research, and gave the first answers on the comparison of two of the four techniques that will be analyzed (triangulation laser scanner - Minolta Vivid 910, and the dense stereo reconstruction tools, PhotoScan). At the conclusion of the project it will probably possible to understand if the cheaper and more portable dense stereo reconstruction technique can give the same results of the more expensive laser scanner technologies. This aspect is one of the most debated in archaeology today. The possibility to record in 3D the site just taking picture can represent a revolutionary change in the discipline, raising an unprecedented dissemination of 3D representation in the archaeological documentation. This discovery can also increase the use of 3D documentation methods between students that want to
understand the discipline and personally test these technologies. So the extreme flexibility and portability of the photogrammetric method can promote teaching, training and learning, giving the possibility to students that cannot have access to the more expensive 3D laser scanner technologies, to experience some of the tools used during the documentation of an archaeological site. In this way they can understand the archaeological documentation process from classes and be trained for the real fieldwork (Di Giuseppantonio Di Franco et al. 2012).

Acknowledgements

The data analysis conducted in this article was possible thanks to the 3D data collection conducted during the archaeological excavation of two sites: Çatalhöyük (Turkey) in June 2010; and Las Cuevas (Belize) in June 2011.

For this reason we would like to thank Maurizio Forte (Professor of World Heritage at the University of California, Merced), director of the Çatalhöyük 3D Dig Project, who allowed us to use the data collected during the fieldwork for educational purposes. Special thanks go also to the Çatalhöyük research project director, Ian Hodder (Dunlevie Family Professor in the Department of Anthropology at Stanford University, and Director of the Stanford Archaeology Center), and field director & project coordinator Shahina Farid (London University College), who allowed testing 3D technologies on site.

Our deepest gratitude goes to the excavation staff of Las Cuevas for the help and assistance during the entire duration of the project. In particular, we want to thank Barbara Voorhies (Research Professor and Professor Emerita at the University of California, Santa Barbara), who supervised the excavation area during the 3D data acquisition, for her collaboration during the units 3D data collection.

We are also grateful to Nicholas Bourgeois (PhD student in World Cultures at the University of California, Merced) and Josue Ramos (Institute of Archaeology – NICH, Belize) for their appreciated help during the data acquisition on fieldwork.

Last, but not least, we would thank Nicolo Dell’Unto (Assistant Professor in the Department of Archaeology and Ancient History at Lund University) and Stefan Lindgren (Research Engineer in the Humanities Lab at Lund University) for their collaboration and help in the brainstorming stage of this project.

References


3D Documentation in Archaeology: Recording Las Cuevas Site, Chiquibul Reserve, Belize
Fabrizio Galeazzi, Holley Moyes and Mark Aldenderfer


Doneus, M., G. Verhoeven, M. Fera, Ch. Briese, M. Kucera, and W. Neubauer. 2011. “From deposit to point cloud – a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations.” In Proceedings of the XXIIIth International Symposium CIPA, Prague 2011, edited by A. Čepek, 81-88. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 6.


Moyes, H., and J. E. Brady. 2012. “Caves as Sacred Space in


Social Spreading of Geometric, Recorded Data from a Range of Types of 3D Scanners via a Web Data Server

Jorge Angas  
University of Zaragoza, Spain

Paula Uribe  
University Bordeaux III, France

Abstract:  
The topic addressed in this paper is the post-processing of the data generated by the scanner and its combination with an UAV (Unmanned aerial vehicle) and its use on a web server. With this aim we have used the project “Documentation, Assessment and Diffusion of Roman Hydraulic Cultural Heritage in the Middle-Valley of the River Ebro (Spain)” as a base. The principal objective of this project was the combination of the trinomial: register, value enhancement and dissemination of the seven different archaeological sites related to Roman Hydraulic Cultural Heritage and the related “water cycle process” – collection, distribution and evacuation- found in Roman architecture. The project was under development from the end of 2010 to February, 2012, and it was directed by the Spin Off of the University of Zaragoza Scanner Patrimonio e Industria, together with the technical advice of the Research Group URBS, which also belongs to University of Zaragoza. The project has also had the support of the Ministry of Culture of Spain and other institutions and the support of the scientist responsible for each one of the archaeological sites.

Keywords:  
3D Scanner, Unmanned Aerial Vehicle, Geometric Data, Dissemination, Web Data Server

1. Introduction  
Currently, the apparent rate of development of technology, together with social immediacy, often does not allow the establishment of an entire process of technologies and procedures necessary in all types of geometric documentation projects. How to focus the latest geometric documentation techniques applied to Cultural Heritage, aimed at a didactic application, is a complex process. The development of this purpose must balance the research process with the social spreading of scientific knowledge. During the last few years, new techniques for geometric documentation applied to Cultural Heritage have been developed.

These techniques have triggered a scientific and specialized revolution in the fields of Architecture, Art and Archaeology. The use of these tools has allowed the precise and detailed reproduction of all kinds of properties and their subsequent morphological analyses. One of these new techniques, which focus on the documentation of Cultural Heritage sites, stands out from the rest by applying a three-dimensional scanning technology (phase-shift, time of flight, and structured light). The information generated by these different methods gives a totally accurate and precise 3D model, and it makes possible to carry out subsequent analyses that will increase the communal knowledge and the documentation of the monument. This technique enables us to implement new methodological focuses through the creation of 3D databases.

However, the lack of a standardized or normalized process in the development and in the management of three-dimensional models, has curtailed the widespread use of available resources. Consequently, in this project the
standardization includes the following three essential premises (Valle, 2007): accessibility, comprehension and geometric utility.

The main contribution to this project has been intended to bring together a methodological consistency throughout the documentation, assessment and diffusion with a website. To do this, we have generated metric documentation, which in many cases did not exist in the fields that make up this project. All this information has been created, filtered and organized from a scientific point of view, allowing continuity in the final process of the transmission of the knowledge, with the inclusion of this information in a web platform.

The results have given an exhaustive topographical record of the principal Roman hydraulic structures in the middle-valley of the river Ebro which guarantees its conservation and structural monitoring.

2. Aim and Scope

The case study is the project called “Roman Hydraulic Cultural Heritage in the middle-valley of the river Ebro”. It has had as a main target the combination of the trinomial: register, value enhancement and dissemination of the different monuments.

The subject matter that lays the foundations of the project is focused on the process called “water cycle” –collection, distribution and evacuation-, aqueducts, dams, sewers, tanks, cisterns and so forth. In order to reach this target, we have selected each one of the archaeological sites according to two main criteria: the architectonical models and the state of conservation. The final aim of the selection of the sites has been to pass on to society, as clearly as possible, the use of hydraulic resources in the Roman period. The applications have been combined on a website by means of different graphic and metric applications.

Our research ultimately aims at carrying out a “democratizing” process of the three-dimensional data. In the same way that a scientific corpus is generated with the obtained data, we believe that it is necessary to spread this data to the whole of society. With this purpose in mind, we have created 3D models of each Roman hydraulic site, on a user-friendly data web server, making a greater dissemination of each monument, and providing a better understanding of archaeological sites. Consequently, this report intends to establish a brief reflection of the state of the enquiry, underlining that one of the present problems is the representation of Cultural Heritage and to
provide a starting point for a potential solution for each one of these problems and thus contribute to the establishment of a coherent development in the process of creativity, innovation, culture and education.

The development of the project can be divided into two parts: firstly, the methodology, and secondly, the case study of the project called “Documentation, assessment and diffusion of Roman Hydraulic Cultural Heritage in the middle-valley of the river Ebro in the north of Spain”. The latter is an excellent example of the process carried out and the integration between scientific data and its dissemination. The same integration process has been followed for both aims. The research has been carried out by the Ministry of Culture of Spain and the University of Zaragoza, working together in search of a standardization of the process of 3D applications in Archaeology.

The different phases established in the project have been divided into:

- Phase I. Previous study and compilation of existing information. Technical advice has been provided by the research group URBS from the University of Zaragoza.
- Phase II. Documentation by means of topographical techniques (3D Scanner, GNSS), terrestrial photogrammetry and low-altitude aerial photogrammetry (UAV).
- Phase III. Post-processing of the information and generation of a database.
- Phase IV. Value enhancement and dissemination of data on a free web site: www.3dscanner.es/Patrimonio_hidraulico_romano.

3. Documentation and Fieldwork Methodology

The fieldwork method involves the use of the different equipment in the different phases, and before starting, it is necessary to establish the aim of the work and the purpose and the type of results which should be obtained.

This new documentation method was created by means of a combination of several techniques: 3D scanning technology (phase-shift, time of flight and structured light) to obtain the geometrical data, the georeferencing of this data by using a GNSS (Global Navigation System).
Satellite System), and the fitting and linking of survey targets with a total station. Simultaneously, terrestrial photogrammetry and an optimization of the texturization of the RGB point cloud was achieved by taking pictures of the monuments with a calibrated metric camera. Finally, an UAV (Unmanned aerial vehicle) has been used for low-altitude aerial photogrammetry.

The technical implementation of the 3D scanner in the documentation field phase arose solely as a link which combined and integrated other documentation techniques such as photogrammetry and topography, forming a single methodological approach.

For the correct choice of technique, there are heterogeneous factors depending on the particular characteristics of each monument documented. Therefore it is necessary to distinguish the different types of instrument used:

• Laser Scanner. Using time of flight equipment model Leica ScanStation and phase-shift equipment model Leica HDS 6100 depending on the site, object distance and immediate environment.

• Structured light scanner (model Artec MHT) to represent the most significant details for each monument, (as the case of epigraphic inscriptions on Muel Dam).

• GNSS-systems. To georeference the location (model Leica Viva GS15), we used the permanent GNSS networks of Aragon, La Rioja and Navarra (recently implemented) through GPRS connection.

Finally, the final composition of the different positioning of the laser scanner was performed using targets (B & W) previously acquired with a total station (model Leica TCRP 1201). The independent creation of a proper coordinate system for each set will allow both the control of final accuracy of the register of each model, and the amplification of and/or monitoring of the documentation area in each of the registered sites in future projects. The application of a 3D scanner allows the generation of an original texture of each architectural structure, and each RGB color value of each of the acquired points is optimised, adding a layer of texture to the three-dimensional model. These textures have been obtained with a calibrated camera (Canon EOS 600D) with different fixed focal lengths - 24mm and 50 mm -. The technique used for the acquisition of the texture at each monument has been different, depending on its morphology, using aids such as poles and other auxiliary elements in Los Bañales Aqueduct (Uncastillo, Zaragoza).

For the further adjustment for radiometric images, we have used a standard 24-color calibration card (X-Rite Color Checker©). The use of this pattern allows the adaptation of radiometric adjustment and the maximization of homogeneity for each of the walls and in the general model obtained. Similarly, it was necessary to store the photographs in RAW format, in which all values are recorded as “digital negatives” without modification and without any image compression. The data is stored in a manner independent of the camera settings.

3.1 Documentation by Means of Topographical Techniques and Terrestrial Photogrammetry

The combination of laser scanning technology and other classical survey techniques is essential, not only for documentation and structural control, but also for the diachronic amplification of the place of study. With this purpose, a connection is established between the generated data by means of a local coordinate system, with geodesic reference to the official datum ETRS89 (Spanish Royal Decree 1071/2007). Subsequently, a survey network is created for the entire work place. This network sets fixed reference points in certain locations around the monument in order to be used as structural control points.
During the first step of the fieldwork, some topographical references (targets) are acquired by means of a total station and a GNSS system, obtaining coordinates of all these references on the same topographic system. Subsequently, a scanner is located at the best positions for acquiring information of the real element, obtaining a large amount of points defining the surface of this element. The resulting data is called a point cloud.

The point cloud requires a special procedure for its treatment and for the improvement of its texture, which is achieved by mapping pictures obtained by calibrated metric cameras, providing a radiometric adjustment to the generated orthophotography.

Consequently, the application of the 3D laser scanning methodology is suggested as an avant-garde technology which is able to solve problems and sidestep limitations which are encountered with the use of other documentation techniques or processes.

In summary, the necessity to apply this new technique for the documentation of these archaeological sites was combined with the difficulty of obtaining fixed information with classical surveying tools and with the possibilities that 3D laser scanning was offering. However, the application of this tool caused other difficulties, related to the management of three-dimensional data. The problems may be summarized as:

1. Computer management of the different formats. It becomes necessary to standardize the format of the data created, in order to obtain generic storage formats like ASCII, which can offer universal compatibility.

2. Graphic representation of the three-dimensional models in two-dimensional electronic media. There is still a long way to go in realizing a conceptual change from the inherited methodological culture.

3.2 Post-processing: Combining the Terrestrial 3D Data with Low-altitude Aerial Photogrammetry

The point clouds acquired by both time of flight and phase-shift scanners were registered in a manner that was supported by the targets (B & W) taken with the total station. The registering and adjustment of the different scans were carried out using different program modules Cyclone© v. 7.3.3 with an average error of alignment between the point clouds of 3 millimeters. The use of each of the permanent GNSS networks in each autonomous community (ARAGEA, RGAN, La Rioja) has
permitted correct georeferencing. This factor will enable the correct location, monitoring and enlargement of the documented area.

The acquisition process of photo textures and RGB point clouds was performed in an independent and dual manner with different objectives:

- Firstly, the photographic acquisition for the mapping of the triangulated model. This was performed with a camera calibrated in RAW format, using a color chart calibrated to perform a radiometric adjustment.

- Similarly, for the acquisition of the color in each scan, given that the phase-shift model Leica HDS 6100 has no camera, and the camera built into the time-of-flight scanner (Leica ScanStation) is very low resolution, we used different settings with a Canon EOS 550 D and a fixed focal length “fisheye” 8 mm. Each of the ball joints and bars are of different metrics in each scanner model, with the purpose of adjusting the tilting axis of the scanner with the center of the camera’s CCD sensor. The end result of this adjustment was used solely for the creation of an integrated web server that integrated all the scans, and was executed using the Internet Explorer® plug-in Leica “Truview”®. Its easy accessibility allows the taking of measurements and coordinate locations through a browser. Working with this type of “embedded web server” allows self-management and the linking of other files (images, databases, bibliographic references, and so forth) to each of the scans, as a local server or directly onto the website. Finally, regarding the incorporation and management of metadata, these files are stored a template in xml format which is editable with any other program (CatMDEdit®, XMLNotepad®), with the raw data from the information capture process.

The triangulation and color adjustment of the model for the generation of the three-dimensional models was performed with 3DReshaper® and Blender®, using the latter for the reconstruction of landscape elements, such as vegetation, lighting and digital terrain models.

### 3.3 UAV Devices

A new evolutionary step in the documentation of Cultural Heritage consists of aerial documentation, by means of an UAV device. This system refers to an Unmanned Aerial Vehicle that can be remotely controlled and its working system can be semi- or fully autonomous, by using a GNSS system. These vehicles, also known as drones, nowadays have numerous applications related to robotics, artificial intelligence, agriculture, natural resource monitoring, aerospace engineering and defence.

The multidisciplinary character found in the variety of UAV models provides this technical achievement, which in one way or another has been able to aid Cultural Heritage. In the last two years we have seen have the arrival of many different types of UAV -some open-source systems- (balloons, zeppelins, helicopters, multicopters, aircraft, etc.) and different manufacturers (Hirobo, Microdron, Mikrokopter, Pegasus and so forth) (Einsenbeiss, “UAV Photogrammetry “, 2009). These arrivals were driven by the advance of low-cost software that provides easy-to-generate three-dimensional models using photogrammetric software (123D Catch®, Photosynth®, PhotoScan®, Ufly®, Pix4UAV®). There is no doubt that in years to come, low-altitude aerial documentation systems will continue to develop along with other technological advances that will make the use of geometric documentation of Cultural Heritage more and more versatile.

The use of this technology consists of the creation of a photogrammetric measurement platform that allows the digitalization of each one of the selected archaeological sites from
aerial images. This platform must operate autonomously by pre-programming the necessary flight paths. This system permits the drone to follow a programmed flight path at each site from geographical coordinates. The incorporation of a high resolution calibrated camera models allows high graphic resolution models to be obtained.

These topographic databases used in the phase of terrestrial registration with the laser scanner provided support for their inclusion in the official datum ETRS89. We emphasize how this gives it a preventative advantage in the future, enabling structural and geomorphological control and also providing the enlargement of the excavation area within the same coordinate system. This factor, coupled with its versatility and low cost when compared with other types of metric documentation tools, makes it a very useful tool in archaeological excavations as it provides quick documentation at each stage of the excavation.

The results correspond to the scientific and informative criteria of the project. Thereby, a series of photographs have been obtained, which by means of photogrammetric and topographic techniques have generated a three-dimensional model and orthophoto of each one of the documented environments, achieving the generation of textured triangulated models and high metric and graphic quality orthophotos. The last chapter in the process of aerial documentation has enabled these results to be utilized also to provide an aerial spherical 360° panorama -view of each and every one of the environments of the archaeological sites in the middle valley of the Ebro river-, with the aim of the dissemination of each and every one of the archaeological sites.

The graphic character this device has provided has been incorporated into the project website under “aerial documentation”. The results have included the generation of spherical aerial photography -utilising aerial planning in the photogrammetric record- which allows the display of each archaeological site (http://www.3dscanner.es/Patrimonio_hidraulico_ romano) from different perspectives.

In each of the flights at the seven archaeological sites, several key points were planned at which the UAV would be positioned at a predetermined position in space. Once positioned at a particular point, the UAV was programmed to perform successive photographic shots every 30°. These spherical panoramas contribute to the whole project, providing a unique graphic document which adds understanding of the landscape from a low-altitude aerial perspective to the dataset. Any user can access these images, which provide
a greater understanding of environment, the extent of the deposit, structures of the features, and so forth. In addition, the metrics generated for each research group consisted of generating orthophotos of each set at different resolutions and three-dimensional models converted GeoPDF and vrml formats, in which measurements can be extracted and sections using free software such as: MeshLab®, Adobe Acrobat®, VRML Viewer®, and Google Sketchup®.

4. Standardization of Processes with the aim of Democratizing the Three-Dimensional Concept

Currently technological innovation has developed much faster than its own understanding and methodological management, and it lacks a corpus of procedures for the correct planning of the process. According to Valle (2007), the standardization of geometrical documentation of Cultural Heritage, should count on, independently of the technique used, three basic premises related to accessibility, comprehensibility and a clear geometric utility. These three guidelines are also perfectly relatable to the three values that integrate and describe the concept of standardization or normalization (unification, simplification and specification). The combination of these has been the theoretical basis of the development of this project, which can be synthesized in:

- **ACCESSIBILITY** through unification. Integration into databases, which are compatible with web servers and may be updated at different levels of access and editing. Unification, which facilitates the integration of the databases accessible and a web platform which is upgradeable and interchangeable with different levels of access and editing. Integration of the metadata set by xml files that are described in the principal information.

- **COMPREHENSION** through the simplification, allowing the use of compatible formats, common and exchangeable with user-level software. Simplification, and with common interchangeable file formats that can mostly be managed with free, user-friendly software i.e. *. GeoPDF (3D), *. U3D, *. Vrml, *. Skp, *. stl, *. dae., contributing to an understanding and therefore easy conceptualization of the dimensional information.

- **GEOMETRIC UTILITY** through the specification, for obtaining measurements and location of elements in absolute or relative coordinates. We are able to check the accuracy of the process, and also it is possible to create a report that any technician is capable of understanding and after being used for checking this accuracy. Specification, through rules that follow the necessary recommendations for the verification of its geometry. The metric component for each project is essential, to be able to obtain measurements and find absolute or relative coordinates, for the capacity of the project to be steered towards infographic alternatives. Fulfilling the scientific binomial informative, always starting from the same geometric file in a logical order of development.

For the standardization of the difficulties in management and dissemination inherent to geometric data generated by laser scanning and UAV devices it is necessary to create applications compatible with user-level software. This allows for simplification in three-dimensional data, apart from achieving the integration of science and popular dissemination. In pursuit of this aim, we are looking for a method for easier management of the registered three-dimensional data belonging to Cultural Heritage, because there have been very few occasions of interaction with other process for registering data, related with other disciplines, such as Engineering, Industry or Environmental Sciences. In Engineering, for instance, it is possible to use the industrial term “as built”
as a metaphorical concept for the recording of reality (Jardi and Angas, 235-242). In these fields there is a higher standardization of the processes that are involved in the control and checking of the quality of the method, which derives from the use of these techniques as principles of authentication of the performed register.

The topic of this research has the final aim of carrying out a “democratizing” process of the three-dimensional result. Thus, in the same way that a scientific corpus is generated with the obtained data, it is necessary to spread this data to the whole of society. With this purpose we have created several documents, in formats that are easy to understand and easy to manage, which result in the greater dissemination of the monument, and also provide a better understanding of the archaeological site for society. This entire process has been managed and supervised by archaeological technicians, as part of a multidisciplinary process (geologists, surveyors, archaeologists and engineers).

Regarding the necessity to democratize, it is possible to define “democratization” in three-dimensional contexts as a multidisciplinary process managed by different technicians.

The methodology used has allowed us to generate standardized applications, integrating the monument in its landscape, as well as studying its architectural morphology. This process has been developed with the use of 3D laser scanning technology, combined and associated with other technologies (surveying, terrestrial and aerial photogrammetry and geodesy). Interest in a standardized scientific analysis lies in the potential of reconstructing and recovering virtually all parts of a building, in its different uses and time periods, and also carrying out a periodic control of the structure, preventing any kind of morphologic pathology.

Other basic enquiries may surround the creation of “democratized” databases, achieved with web servers that are easily configured, aiming at a greater dissemination of the archaeological site, graphically, but principally, in a geometric way. This allows for the quick and accurate display of any element or part of the structure.

5. Results and Conclusions: Dissemination of Cultural Heritage by Means of Applications Obtained From 3D Scanning and UAV Data

The development direction of the post-processing method has as a final aim the “democratization” of the three-dimensional results in each one of the cultural sites documented (Roecker, 341-355). With this purpose, several documents can be created in
formats that are easy to understand and easy to manage, by scientists and the whole of society, and it is resulting in a greater dissemination of the monument and in a better understanding of the archaeological site by society.

The adaptation of the results for a web page is designed as a continuation of the same methodological process. To do this, the results have been adapted to formats compatible with most software at the user level, which often require compression and reduction of their graphical quality for the benefit of social diffusion in a web environment.

Transforming files to compatible formats have been established mainly in:

- Graphical and metric formats.
- Playing a selection of scans using html formats.
- Sets of planes of each set: plans, elevations and sections.
- Graphical formats. Derived from metric applications.
- Videos with planned itineraries within the 3D model with links to a selection of spherical panoramas captured by the scanner. The originality has been to offer a real three-dimensional view of the cultural sites while avoiding virtual recreations, developing documentation which can be valid to a scientific and to the whole society.
- Spherical images of the flight planning carried out by UAV.
- Additional graphical information.
- Location of each set and images of each of the monuments in the aerial and terrestrial documentation process.
- Didactic information.
- Chapter on the methodology used in each stage, historical introduction, glossary of architectural and historical terms, videos presented by the site scientists, teacher’s book with activities for high school students, literature and various web links.

Figure 6. UAV flight in the Roman dam of Muel. The 3D mesh was obtained with the combination of laser scanning data and photos taken at different positions with the calibrated camera of the UAV.
Due to the heterogeneity of each one of the documented cultural sites, several points require special attention:

1. Registration of each archaeological site in its current state, regardless, if necessary, of other external elements or modern architecture - as an example in the sewer of Caesaraugusta- in order to understand the relationship between Roman architecture and its natural environment. For the registration phase, methodological criteria have been established with other processes related to exogenous disciplines such as engineering, industry and environmental sciences. All this involves a new language that allows us to register reality; to establish points of contact with other sectors contributing to the acquisition of an overview on how to organize and manage the acquired three-dimensional information. Therefore, we believe that the exchange of working methodologies with other disciplines with the aim of spreading Cultural Heritage sites will be one of the objectives in the coming years, particularly in the fields of management and dissemination of the information.

2. Standardization of processes. Quality control and tests by means of standard procedures, ensuring interoperability and communication of information through unification, specification and simplification. This facilitates the understanding of the different processes for reaching the final result, with it being possible to analyze individually each one of them.

3. Process of “democratization” of 3D results, that provide spreading and dissemination of the information, by using compatible formats and open-source software, which are free and easy to manage. By this process it is possible to arrive at a minimum level of access and understanding of the information to the rest of society.

4. Interdisciplinary methodology in the process of the study coordinated by a technician, in this case an archaeologist. The basic purpose is to facilitate management through easily configurable web environments with different degrees of access, in order to achieve greater dissemination of the archaeological site, graphically and especially geometrically. This allows, in a fast and accurate way, the viewing of any element or part of its structure. Thus, there has been a combined development, valid for a greater appreciation and spreading of the information obtained. Apart from contributing to transdisciplinary research itself, and bridging the different methodological and dimensional conceptualization gaps, it contributes to an integrated point of view that avoids the segmentation of the information channel. With this aim, different topographical techniques have been combined with different 3D scanning technologies, always depending on the particularities of each one of the sites documented.

In summary, the results of the project have been:

- To obtain three-dimensional models of different Roman hydraulic typologies, improving on previous documentation.
- To use the same information for scientific and non-specialized public, offering the project on a public access web site: www.3dscanner.es/Patrimonio_hidraulico_romano.

Developed as a flexible environment where it is possible to add different kinds of information. This concept is part of a database with two clear functions for the future:

- To increase the documentation obtained for each monument.
- To monitor any component in its structure.
- To document each structure in the current
state of conservation, making a real restitution of the cultural site. In this way it is possible to offer an explanatory realistic view of the site and especially of the space that it occupies in the archaeological landscape.

- Regardless of this project, the most important factor is that all this metric information acquired is stored in a “documentary database”.

We have managed to combine a greater diffusion and appreciation of the obtained information, with a contribution towards multidisciplinary scientific studies, thus solving the various existing gaps in three-dimensional methodology and conceptualization applied to the study of Roman Hydraulics.

I would like to emphasize that it is necessary to invest in the creation of applications that are compatible with user-level software. This would simplify the problems related to three-dimensional data sharing and expand possibilities for working.

There remains, however, much to explore, and with this project we have only pretended to approach a new way of managing recorded information. Therefore, we understand that the exchange of working methodologies at the limits of various disciplines, for dissemination of Cultural Heritage, will be one of the objectives in the coming years, especially in terms of information management.

**Acknowledgments**

This project has been financed by the Ministry of Culture of Spain. We also would like to express our gratitude for the cooperation and kindness of the scientific staff of all institutions that have collaborated in its development: Research Group URBS, Government of Navarra, Government of La Rioja, Government of Aragon, Council of Zaragoza, Museum of Teruel and the Uncastillo Foundation.
References


Combining Terrestrial Laser Scanning and Techniques of Digital Image Processing in “Archaeology of the Architecture” Analysis in the Walls of the Andalusian Site of Vascos (Navalmoralejo, Toledo-Spain)

 María J. Iniesto-Alba  
 University of Santiago de Compostela, Spain

 Miguel A. Bru Castro  
 University Autónoma of Madrid, Spain

 Estela Paradelo Fernández and Pablo Carballo Cruz  
 University of Santiago de Compostela, Spain

Abstract  
This paper presents the work of 3D recording and geometric documentation of part of the walls of “Ciudad de Vascos” through specific methodologies of Archaeology of the Architecture, with the objective to obtain ‘quickly’ the individual measures of every element which compose the wall. We carried out a topographical survey with a TLS (LeicaScanStation C10), for the geometric and volumetric reconstruction; in addition we used a GPS to give absolute coordinates to the work. We obtained the 3D textured model and a series of plans or sections, like first results of the project, as well as a series of orthoimages that will support to the archaeological study. Finally, we used techniques of digital image processing, mainly: image enhancement, edge detection and filters, in order to ‘automate’ the process of defining structural elements to obtain the measurements requested.

Keywords:  
Terrestrial Laser Scanning, Image Enhancement, Edge Detection, Archaeology of the Architecture, “Ciudad de Vascos”

1. Introduction  
The aim of this paper is to present the work which we are developing in 3D recording and geometric documentation of part of the walls of the Archaeological Site of “Ciudad de Vascos” in Toledo – Spain; to achieve one of the objectives of Archeologia della Architettura the chronology about the typology of the masonry of the buildings (Francovich and Parenti 1988; Quiros, 2002). To delve into this subject, we were looking for a geometric definition and measurement of all elements which compose the wall, and to arrive at this propose we support the study with a survey with Terrestrial Laser Scanner (TLS) technology with the implementation of this data with techniques of digital image treatment and edge detection (Like in Lambers et al. 2007/ or Natividad and Calvo 2010).

Basically, we propose the use of orthophotos generated from the processing point cloud obtained with a TLS. In these images, we applied techniques to enhance and improve, looking to extract the edges and to define every element of the structure. The objective of this process has been to facilitate the extraction of numerical data of all the elements of the walls, its dimensions (length, width, area, etc.). For that reason we suggest the implementation of programs used today by the biological sciences, to counting and measuring cells, thus contributing in the mensiochronology methodology that we try to apply at the study.

Corresponding author: mariaj.iniesto@usc.es

376
Combining Terrestrial Laser Scanning and Techniques of Digital Image Processing

María J. Iniesto-Alba et al.

This brings ‘quickly’ both data, the definition of the geometry and the numerical data to manage with statistical programs. In consequence, these studies, trying to go into the methodology of ‘Archeologia della Architettura’, contributing with this idea to arrive at a process of automation of the buildings study.

2. Contributions Between Geomatics Techniques and Archaeology of the Architecture in this Particular Archaeological Site

To study cultural heritage elements using by Archaeology of Architecture method, the stratigraphic readings of the elevations and their type must be analysed. Firstly data are collected, in a critical and analytical way, and secondly the material remains have a stratigraphic study, in order to translate them into suitable documents to be interpreted historically. The stratigraphic reading of the elevations may be organized in three phases (Quirós Castillo 2006):

- I: Definition of archaeological Stratigraphic Units. Drawing and representation of any Stratigraphic Unit and its peculiarities.
- II: Absolute dating of stratigraphic sequence, according to structural or temporal arguments into phases and periods of the stratigraphic records.
- III: Finally, the subsequent analysis of the process in the archaeological site is made, that is the historical interpretation of the sequence and the project definition (the typological study).

Without going into great detail, our work is focused specifically on Phase I, which contains the delineation of the structural elements, for which some method of data collection and processing of spatial data is needed, finally we could obtain the 2D and/or 3D representation of the study element; and Phase II, in which measurements of dimensional features of the elements are made as the study base as archaeometric methods, in particular for the mensiochronology technique (Quiros 1996) (Fig. 1).

In this particular archaeological site, we have known through the different excavations, alongside specific analytical tools applied to the wall of the city, that this madīna -city- «called Ciudad de Vascos» could have had an occupation between the IX and XI centuries (Izquierdo 2005).

That information is very important because that period was the control of the Umayyad dynasty, from Córdoba, when the important wall of the madīna was erected. Hence, the wall-fortress was erected alongside...
some constructive parameters which include common features founded in the Umayyad architecture, either in Middle East or in the Iberian Peninsula. In any case, some differences within the same building allow us to define probable chronological or constructive phases, allowing us to discuss about local variations, which we should find in every dimension of the building.

Our objective is going into the first and second phase of the methodology, to serialize the building functionally and chronologically through the measurement beyond the data. Therefore, one of the main elements to take into account is the analysis of the variation and similarities of the modules and measures of the edification (i.e., dimensions of ashlar masonry...).

Specifically this study is located on the west side of the wall, between the first and second towers, and the one of the main entrances to the city, the West Gate, a special point of political propaganda as demonstrated by the horseshoe arch carved into the ashlar. In these areas of the wall we can see a sample of the ‘official’ construction, which have interpreted like an Umayyad building (Torres 1957; Pavón 1987; 1991; Zozaya 2009).

3. Materials and Methods

The methodology detailed below has been divided both in fieldwork and office work. The necessary fieldworks for the empirical study are composed mainly of archaeology and topographic surveys. The office work has been more diverse and may be mentioned from the literature review, the data processing and coordinates calculation, the digital image processing and the dimensional measurements of ashlar and other structural elements that form the studied section.

3.1 Field Data Collection

We carried out a 3D T.L.S. survey with a Leica Geosystems ScanStation C10, on the west side of the wall of “Vascos”. In total, there were 12 T.L.S. locations, taking one point every 2 mm, of which seven (V-1 to V-7) are outside the wall, and the rest in the interior (V-8 to V-12). The results of these scans are a point cloud of 13,647,654 points, formed by the coordinates (X, Y, Z), the three colour components (RGB) and the intensity (I) of each record point of the wall (Fig. 3). The planning for this phase is absolutely necessary as the optimal location of the scanner ensures the complete definition
of the element, due to the complex topography and the possibility of occlusions and shadows. Targets were used to georeferencing work, positioned so these may be visible from several scans, to which they were given absolute coordinates using GPS techniques and the RTK method.

3.2 Registration

The office works began processing the data captured in the field with the T.L.S. The first step was the registration of multiple scans that consist of aligns georeference and combines the multiple data sets in order to get a complete 3D cloud. It was used on the module Cyclone Register.

In our case, the registration was done by the software, in an iterative process for which at least two common points in each point cloud are necessary, these are the registered targets. In this case it may be complex because the masonry is not defined by clear edges, but they are rounded, and there are no clearly defined points that can be identified without doubt in the different scans.

3.3 Cleaning the Point Cloud

At this stage, we cleaned all the noise of the scans, mainly undergrowth and vegetation, leaving only the wall remains. This information to delete was completely removed without possibility of recovery, and it made completely in manual form (Fig. 4).

3.4 Meshing

Geomagic Studio 12.0 software was used to obtain the triangles mesh. This software has a large computational power, and supports a comprehensive range of 3D scanners in XYZ / ASCII format, like the Leica data in .ptx file format, which includes the coordinates, the colour information (RGB) and intensity corresponding of each point. But in this case, it was decided to transform the data format to .pts file format (only coordinates) and let the default appearance that the program gave to the mesh (Fig. 5).

Furthermore, due to the large amount of data captured, a single mesh for the area of study becomes very heavy, it was necessary to divide the area of study into smaller parts thus to determine different zones and subzones.

The main problems at this stage were that all zones had hidden areas by the vegetation.
and undergrowth although the wall has been cleaned before the survey; besides there were trees that we could not cut off for Conservation Nature reasons.

3.5 Orthophotos

With the same software (Geomagic Studio 12.0) textured ortho-images of all elevations were generated. Which was a product derived from the triangle mesh with a resolution of 96 pixels per inch, whose aims was to provide us the basis for future treatments, like the archaeological study of the structural elements (Fig.6).

In this case, we decided not to work with photographs, which came from the textured model, our objective was to subject these ortho-images of different digital image processing.

3.6 Line Drawing of Structural Elements

This process can be accomplished directly from the 3D point cloud, i.e. using the point cloud as base from which geometric features are traced, thus creating a vector model, or delineating manually from orthoimage. But our proposal is to use image treatment and enhancement techniques, to extract the edges of the structural elements and ‘automate’ the process.

The software that we used in this part of the work was GYMP 2.6., which has the added advantage of being free software.

The main aim was to highlight the edges of each one of the elements that form the walls. They should be treated as a separate object. In an image, which defines that an element is distinguished from another is the difference between Digital Levels (DL) of each one, that is, each image pixel has a lower or higher DL than their neighbours. Thus, an element is defined by a set of pixels adjacent with identical or very similar DL value, and its limits are marked by a change in the DL value. By spatial contrast, i.e. the difference between the digital level of a pixel and the neighbouring pixels, we define each object using filtering techniques, smoothing or enhancing these contrasts, and edge detection.

Firstly, a phase of noise reduction is performed, consisting in removing residual information by the application of the ‘Gaussian filter’ and a ‘Non-linear filter’. Then, edges were delineated by the application of the ‘Sobel’ and ‘Laplacian’ filters. The results were not completely satisfactory, since the ashlers and masonry were not entirely demarcated, for that we looking for an alternative. ‘Tampon filter’ of Photoshop simplifies the image to appear stamped with a rubber stamp, which shows the pixels continuously, creating closed shapes more or less homogeneous and resulting in the best way to define the ashlers, and other elements of the wall, as individual and closed forms.

3.7 Dimensional Measurement of Structural Elements

The last phase of our work was to calculate the dimensional measurement of each of the structural elements, mainly the ashlers and masonry areas. The Aphelion® Dev software has been developed by the biologists to carry out counting and measuring cells for further studies. In that direction we though to use this program to determine the width and height of
each element. One of the requirements of the program is to work only with binary images, i.e. black and white, so it was necessary to divide the process into three phases: defining a reference standard and setting the measurement units (inches, centimetres, pixels ...), excluding values for unwanted digital levels, since images resulting from treatment with Tampon filter because they are not exactly binaries and finally launching in the program the measure of the elements of study. Thus, it was possible to detect defined surfaces (closed) with the same digital level (Fig. 8).

As the result, a table with the number of ashlar or masonry element and the height and width is obtained, necessary parameters to proceed to the temporal study by mensiochronology.

3.8 Results Verification

Finally, the process of results verification helps to validate the obtained results. Thirty elements for each study zone were randomly chosen, its measurements were taken in site, the height and width of the ashlar and masonry stones that make up each facade, carrying out a statistical analysis that consists of calculating the residuals and the study of the error behaviour for each study zone, to give the quality value of our work.

To set a maximum error to help us in defining the accuracy and precision of the work, we follow the same criteria of any topographical survey, where accuracy is given by the limit of visual perception (0.2 mm * N, where N is the denominator of the scale). Whereas we work with a representation of the information to 1:100 scale, this value would be two centimetres. On this basis, we can calculate what percentage of the residuals of each sample is among the values of -2 and +2 centimetres. This criterion was fixed between both parts of the team. The archaeological reason has been taken in 2 centimetres, because it is possible to obtain a medium-range of the ‘official’ measures, which the andalusies constructors may be used, and the difference could come for the limitation of the construction tools.

On the other hand, to analyze possible systematic error sources, a graphic representation on the study zones was made where, in a visually way, through a colour scale it can be seen how residuals are distributed.
4. Analysis and Results

Even though we have already focused to the information about how the analysis of graphic data has been accomplished, and how the measure of each element was calculated. We would like to present some of the main problems we came across when trying to define different elements with that method. Afterwards, we are presenting a brief abstract of the analysis of tolerant value.

Regarding the delineation of structural elements, the mainly faced problems can be summarized in four points:

• Define problems related with the definition of irregular stone mason edges.
• Errors caused for the joints of lime mortar.
• Other problems have been faced regarding certain singular elements, such as waste pipes, which had to be defined manually.
• Finally, an important aspect to deal with was the hidden areas in the digitations. Those areas had to be completed with photogrammetry or another topographic system.

Regarding the general study of tolerance, as mentioned previously, while taking measures, the team chose to do a sampling in each of the four areas studied. Such sampling consisted of taking a total of thirty measures within the site, to contrast them in the laboratory and see the tolerance of the tool in a scale 1:100. Such tolerance was of 2 centimetres, and two variables could be founded: length and height, both of them will be analysed as follow.
On the first table, subareas 1 and 4.1 can be detected. In the first one, we can observe that the better results lie in the length of the measures, with a ninety one per cent (91%), while the heights have an index higher than the tolerated one. Studying the probable cause, it is possible to detect that the higher output of tolerance can be founded in the irregular stone masonry. Furthermore, it can be observed in that table the surpluses of the length. In only few cases they exceed the two centimetres. Regarding the measures in 4.1, an average index of output of tolerance was detected without a clear pattern of distribution.

In the second table, the subareas 3 and 7 can be observed and also how the success of the tolerance is higher. On the third subarea the height has most of the residuals fluctuating between 0-2 cm. Two cases lie in 3 cm, which represents the ten percent of the total.

Finally, the subarea seven presents some output of tolerance, probably due to the irregular masonry it can be spotted how the length caused an output of tolerance higher in the data. That can be due to the masonry and the need for improvement in the leak of the images. In this case, we wanted to present a

![Figure 10. Length remainders distribution in centimetres, Subzone 1.](image)

### Table 1. Subareas 1 and 4.1, wall of madina Vascos.

<table>
<thead>
<tr>
<th>SUBAREA</th>
<th>PERCENTAGE OF VALUE IN TOLERANCE</th>
<th>DISTRIBUTION OF RESIDUALS</th>
<th>ORIGIN OF RESIDUALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LENGTH 91%</td>
<td>Irregular stone masonry, the highest residuals</td>
<td>The irregular stone masonry affects the delineation with image filters</td>
</tr>
<tr>
<td></td>
<td>HEIGTH 57%</td>
<td>Irregular stone masonry, the highest residuals</td>
<td>The irregular stone masonry affects the delineation with image filters</td>
</tr>
<tr>
<td>4.1</td>
<td>LENGTH 56%</td>
<td>Uniform distribution</td>
<td>There is no clear pattern</td>
</tr>
<tr>
<td></td>
<td>HEIGTH 77%</td>
<td>Uniform distribution</td>
<td>There is no clear pattern</td>
</tr>
<tr>
<td>3</td>
<td>LENGTH 84%</td>
<td>Higher residuals from the centre to the left</td>
<td>Decreasing cloud of points affects the photo rectification</td>
</tr>
<tr>
<td></td>
<td>HEIGTH 90%</td>
<td>Most residuals fluctuate between 0-2 cm. Two cases lie in 3 cm.</td>
<td>No uniform pattern</td>
</tr>
</tbody>
</table>
visual analysis of the percentage in the output of tolerance.

With this study we would like to present some preliminary conclusions, assuming that it is necessary to debug the job related to the post-process of images, even though we can observe that in a scale 1:100 there is an average of 70% of the data (including irregular stone masonry), due to obtain a medium-range.

5. Conclusions

Making a comprehensive assessment of all the work, we could conclude that the data collection by T.L.S. has allowed a greater speed data capture, resulting in significant reduction of time-consuming in the fieldwork and thus, in a reduction of costs, providing an accurate and high detailed three-dimensional record of the walls, impossible to achieve in time and cost by traditional methods (Peripimeno 2005).

One of our main aims was to be able of simplify all processes, although human intervention is required in some specific tasks (i.e. cleaning of some remains of shadows of the point cloud), the phase of the line drawing could be completely automated by the application of edge detection filters and techniques of digital image processing. This graphic record supplemented with the software application for extraction of the dimensional measures of the elements to be studied, allows us to obtain in a few seconds measurements that, performed by hand, would result in a high temporal cost, depending on sample size.

This methodology developed between archaeologist and topographers, arrives to an important result for the study of the historical buildings, and it has the ability to perform analysis quickly, easily and reliably for a comprehensive historical study of the chronotypology. Future lines of work should further deepen the application of these techniques, which allow the possibility of turning raster information in vector for the subsequent application of the same in, for example, a GIS environment.

By improving this method applied to historic buildings we can achieve the complicated objective of obtaining mensiochronology of any stratigraphic unit and any individual typology. That will enable in the future a quick comparison among buildings chronologically and typologically similar, allowing us to go further in our understanding of the techniques and possible ways to construct in the Middle Ages.

References


1. Introduction

Tangible cultural heritage documentation can be approached from various angles because of its diversity. Yet, all of them agree on the need for geometric documentation as the basis for the knowledge, conservation, restoration or simply dissemination of the monumental remains of past cultures.

The geometric documentation of the elements of cultural heritage, either archaeological sites or historical buildings, refers to both the shape and the size of the elements and their spatial arrangement from a local and global perspective. Geometric documentation can be 2D or 3D and ranges from maps at different scales, orthophotos, floor plans, elevations and sections, to perspectives, 3D models or virtual reconstructions.

Today, the different phases of the geometric documentation process, i.e. data capture, processing and output, are undergoing a vertiginous process of change. In recent years, computers and technology have remarkably evolved and have provided geometric documentation with a variety of means that have brought about new procedures and methods. Such procedures and methods tend to improve data capture systems and process automation, thus increasing productivity and improving the quality of results.

From the point of view of data capture, various techniques are used. Currently, traditional techniques such as topography, GPS and conventional terrestrial and aerial photogrammetry coexist with more modern techniques, such as laser scanners, both terrestrial (TLS) and aerial (LIDAR), or the most recent Unmanned Aerial Vehicle (UAV) photogrammetry.

Deciding on the most suitable technique for a given project will depend on the aims and the available resources. However, there seems to be no single answer. In our opinion, the solution is not excluding but combining and integrating different techniques, particularly in cultural
heritage applications, in which combining the best geometry of some techniques and the best radiometry of other techniques seems an interesting line of action.

The research presented in this paper integrates UAVs and TLS technologies without neglecting other more traditional techniques such as GPS or digital terrestrial photogrammetry for the complete topographical survey of the Roman walls of Lugo with high definition in some representative zones, such as the Miñá gate. The result was a 3D model of the Roman Walls, plus a series of products, such as plans and orthophotos, which can be used to provide extremely precise measurements for the study of the geometry of the walls and the analysis of their structural problems and distortions, particularly in areas that have suffered a greater degree of degradation.

2. Background

2.1 Terrestrial Laser Scanner (TLS)

Early in the 21st century, 3D laser scanning revolutionized the automated capture of large point clouds of an object’s surface in a systematic pattern at a high rate (hundreds or thousands of points per second), achieving the results (i.e. 3D coordinates) in (nearly) real time and with associated intensity or colour values (Böhler 2006).

Today, TLS is a fairly mature technique with multiple applications related to cultural heritage, mapping, civil engineering (structural control and inventory), industry, environmental studies, forestry and geology, medicine (prosthetic design and dental surgery), criminology (accident reconstruction and interior scenes) and tourism (virtual tours). Yet, research in this technology is still open, but in process of consolidation and exploitation.

During the last few years, TLS has been a great field of research and advances in measurement and representation techniques for the geometric documentation of heritage. Actually, there is an extensive literature on a variety of heritage structures and elements, from small pieces and findings to large monumental, isolated buildings and archaeological sites, both at the national (Biosca et al. 2007; Buill and Nuñez 2008; Mañana et al. 2009; Baceiredo and Baceiredo 2010; Lerma et al. 2010) and international level (Ahmon 2004; Peripimeno 2005; Arayici 2007; Lambers et al. 2007; Al-Khedera et al. 2009; Entwisle et al. 2009; Ruther et al. 2009 Schueremans and Van Genechten 2009). This does not mean that this is the only optimal technology, but it must be taken into account, mainly if there are no automated photogrammetric procedures from multiple images available (Petti et al. 2008).

Besides being a non-invasive, light-independent technique that captures data fast and can be used to digitize any element or terrain, TLS has the advantage of providing immediate measurements with great precision (even millimetres), which allows for immediate digital reconstruction and 3D modelling from point clouds. Moreover, TLS is an excellent sensor for the documentation of irregular surfaces or complex geometries such as altarpieces, arches, sculptures or archaeological sites and buildings with great ornamentation (sculpture, columns, pilasters, medallions).

However, despite their advantages, laser scanning sensors have also some drawbacks, e.g. the technology is still expensive, it does not allow for the selection of a specific point, and random point selection often generates imprecision, especially in the exact representation of object edges. In addition, although most TLS incorporate image sensors that provide radiometric information, the resolution of radiometric information is low. Some post-processing operations such as point cloud clearing are time-consuming and laborious. On the other hand, the development of methods and tools for the automatic acquisition and efficient extraction of the geometry of different
elements in order to obtain a final product suitable for heritage documentation, remains an open line of research.

Therefore, laser scanning is not a comprehensive and definitive solution for 3D modelling, at least in applications related to different modes of architectural and archaeological heritage (Böehler 2006; Guarnieri et al. 2004; 2006; González Aguilera 2007; El-Hakim 2007; Pavlidis et al. 2007 and Guidi 2009) in which modelling requirements and complexity significantly increase from the point of view of geometry as shape (González Aguilera 2007).

2.2 Unmanned Aerial Vehicle (UAV)

An Unmanned Aerial Vehicle (UAV) is an aircraft without a human pilot on board. The United States Air Force (USAF) started investigating the use of UAVs in the early 1960s, but civilian research did not begin until the early 1990s. With a growing number of civil applications, because of their potential for remote data acquisition in dangerous environments and/or inaccessible zones, rapid and low cost, UAVs are gaining more and more attention, both for scientific and commercial civilian purposes. Agriculture, environment, engineering and civil protection, among other areas, have successfully adopted this technology for many of their tasks.

From the perspective of spatial data capture, UAVs equipped with a digital camera have revolutionized the world of photogrammetry. UAV photogrammetry can be understood as a new photogrammetric measurement tool (Eisenbeiss 2011), a low-cost alternative to traditional aerial photogrammetry systems on manned aircrafts that can reduce both the cost and time of conventional flights. In addition, UAVs fly at low altitude thereby increasing the resolution of the captured images and opening up new applications in the domain of short distance and the combination of aerial and terrestrial photogrammetry.

An UAV system for photogrammetry consists of an aerial based platform (helicopter and airplane, principally) equipped with imaging cameras, including video cameras, thermal or infrared camera systems, multispectral cameras, range cameras sensors and airborne lidar sensors, or a combination thereof depending on the payload of the UAV. Furthermore, for the determination of the trajectory, UAVs feature by default an integrated GNSS/INS system (global navigation satellite system/ inertial navigation system), barometric altimeter and compass systems (Eisenbeiss, 2011).

Only a few authors have been concerned with UAV systems used for the acquisition of aerial photographs in cultural heritage because this is a very recent technology. Yet, there is growing interest towards this technology because of the aforementioned advantages. UAV photogrammetry has been used to data capture in the archaeological site of Pinchango Alto, Peru (Eisenbeiss et al. 2005); Copán, Honduras (Eisenbeiss et al. 2010); Vergina-Aegeae, Greece (Patias et al. 2007); Drapham Dzong, Bhutan (Gruen et al, 2009); and Pava and Veio, Italy (Sarazzi and Taufer 2008; Chiabrando et al. 2011), with satisfactory results that offer a promising future for the use of these systems.

3. The Study Site: The Roman Walls of Lugo

The city of Lugo, Galicia, Spain, is built on the remains of the old Lucus Augusti, surrounded by Roman walls (Fig. 1). The walls were built in the later part of the 3rd century by the architect Vitruvius and are located in the historical city centre. The Roman walls of Lugo are among the best-preserved walls in the world and are the only ones of their kind with intact walls. The entire length of the walls is about 2 km, enclosing an area of 34 ha. In 2000, the walls were inscribed on UNESCO’s World Heritage List as ‘the finest example of late
Roman fortifications in western Europe, and have also held the status of Spanish monument since 1921.

The Roman Walls of Lugo consist of 70 towers, spaced at regular intervals around the walls, and 10 gates (Fig. 2), five of which are ancient and date to Roman times. The other five are recent and were built after 1853. There are five stairways and a ramp, which give access to the parapet walk, as well as a number of double staircases within the thickness of the walls, which give access to the towers from the parapet (Fig. 3).

4. Materials and Methods

In our research, we carried out a survey of the walls' boundary including the internal and external stone facings and the parapet. We used two different technologies for data capture (Fig. 4), a TLS and a UAV system. Yet, we needed other technologies such as GPS, which was necessary to provide support for data capture and to georeference our data in a global coordinate system.

Thus, the methodology can be divided into four fundamental phases: establishment of a micro-geodetic network around the walls; TLS data collection and processing; UAV image acquisition and processing, which include the phases of planning, capture and processing; and, finally, combination of TLS-UAV data and generation of results.
4.1 Micro-geodetic Network around the Walls

The shape and size of the wall (around 2 km) required the use of multiple TLS stations and several working days. In addition, in certain areas, future actions have been planned that require the use of conventional surveying techniques. To guide and carry out the work, and to be able to link directly the data for future actions, we need a micro-geodetic network around the wall and a series of fixed stations secured by Allen screws, where any type of topographic instrument or targets can be fastened (Fig. 5).

Because a geodetic link was required to provide network stations with absolute coordinates, a reference station was placed on top of the SIT Building (Spatial Information System at the University of Santiago de Compostela), and its readings were taken from three REGENTE (Red GEodésica Nacional por Técnicas Espaciales) geodetic points by GPS techniques, using the method of static positioning. To give coordinates to the station, the coordinates were measured using GPS techniques, particularly the Real Time Kinematic (RTK) method. In this work, we used Trimble5800 GPS receivers.

4.2 TLS Data Collection and Processing

In the planning phase, we studied the optimal location for the laser scanner in terms of data collection, considering the characteristics of the equipment, such as sensor-to-object distance or occlusions, and the difficulties of the area, among which flow of cars, vegetation and presence of a large number of people, at particular times of the day, insofar as the walls are a leisure area and the commercial and cultural centre of the city. The planning phase is essential to ensure a complete survey of all elements.

For point cloud collection, the Trimble GX Advanced 3D laser scanning system (www.trimble.es) was used. The scanner features for the field of view are 360° in the horizontal direction and 60° in the vertical, both 38° up the horizontal and 22° down the horizontal, which is an important factor that can limit fieldwork. To ensure the required quality, the fast-mode scanning method was used by setting mesh resolution at 2x2 cm. The longest-range scan mode was used in specific areas where there were details or elements of interest, or where further actions were planned, as in the Miñá gate, where a study of deformation was carried out. The mesh resolution defined for the longest-range scan mode was 1x1 cm.
The placements of the scanner were connected to ensure proper orientation between the different point clouds. In this case, a series of Trimble-brand targets were used.

To improve the quality of the images, we captured digital images with a Canon EOS 400D camera with 10.1 megapixels and a focal fixed at 18 mm, in specific areas with details of interest because the camera of the scanner did not provide the desired quality for this work. The photos were taken using terrestrial photogrammetric processing.

Once the fieldwork was completed, point cloud processing was carried out in the office. The acquired 3D points were processed using the Real Works 6.32 software by Trimble, which is specifically designed to process data from laser scanner surveys.

The first procedure in the processing of scanned data consisted of registration and georeferencing operations, which was accomplished by matching common targets. The next step was data filtering and cleaning in order to eliminate the noise from the image (Fig. 6), alignment of partial scans, and hole-filling. Then a mesh of triangles was created by using the 3D mesh method to transform the point clouds into a triangular mesh. We generated multiple meshes of the different parts of the wall because a single mesh was unmanageable for a PC (Fig. 7). Following the generation of the meshes of triangles, a simplification of the polygon mesh was required to optimize model management by the computer. The last procedure in this phase was texture map processing in some areas. The photographs must have been previously processed with digital image treatment software, PhotoShop CS4 in this case, to balance the hue, brightness and contrast of the images. To georeference the images, homologous points between the images and the point cloud were used. More than five points were used for each image. Finally, the process automatically allocated texture. To achieve that, the images were combined with the mesh of triangles, which produced the real 3D model of some areas of the wall, such as the gates known as “Santiago” and “Miñá”, plus a series of products for specific studies and analyses of the wall (Fig. 8).

4.3 UAV Image Acquisition and Processing

The parapet images were obtained with a Microdrone GmbH MD4-200 UAV (Fig. 9) because of the difficulty of obtaining them with other systems such as helium balloons or complex systems of poles. The camera of the system was a Pentax Optio A40. The drone
was a helicopter with four propellers that contributed to a very stable aerial vehicle. Since the drone was fully equipped with sensors such as GPS and INS, it could fly, take-off and land autonomously.

Before take-off, a flight plan must be developed by defining the area that must be flown over on a map, based on Google Earth. The microdrone software automatically generated the flight plan based on a number of parameters, among which altitude, overlap or camera. Image acquisition was performed according to the flight plan.

In this case, we flew only over small areas of the wall to improve the 3D models of some areas of interest and to assess the possibility of flying around the entire wall, which would require special permits in order to close the walls during image capture (Fig. 10).

The next step was image processing. We considered two options:

• processing images using the conventional photogrammetric method with non-metric cameras, which involves camera calibration, image pre-processing, aerial triangulation, digital terrain model (DTM) extraction and orthophoto production, or

• performing a simple image rectification by using a DTM from the TLS process to get the orthophoto.

As stated above, our first goal was to get the textured model of the parapet in some specific areas and to generate the orthophotos from the textured model as a derivate. In this case, we have the cloud of points from the TLS and, consequently, the DTM required for orthophoto generation and image rectification. Accordingly, we chose the second option, such that we only had to pre-process the images to do the radiometric adjustment to improve image quality and, finally, establish the correspondence between the image points and the model points in order to obtain the rectified images from the UAV and the textured model.

The radiometric adjustment of the UAV images was made with PhotoShop CS4 also in this case. The processing of TLS data was performed as described in section 4.2, with the only difference that the process of data filtering and cleaning was reduced, as we found less noise in the point cloud of the parapet, and a single triangle mesh was generated for the areas of interest. Finally, homologous points were used to georeference the images and to process the texture map in these areas. At least five common points in each of the images and the point cloud were selected.
5. Results

The digital 3D model of the point cloud and the meshed surface of the Roman walls of Lugo are the first results of this work. Both outputs have been obtained directly from the TLS. The raw and recorded coloured point clouds generate a complete, comprehensive and continuous model, suitable for use in various applications. Overall, the estimated accuracy of 3D point coordinates is less than 2 cm (Fig. 11).

Line drawings are generated from the mesh of points by using the point cloud as the basis from which geometric features are traced, thus creating a vector model (Fig. 12) and a number of plans, sections and elevations at 1:50 scale (Fig. 13- right).

Another output is the textured surface model of some areas of interest, which has been obtained by combining TLS data and terrestrial and UAV images (Fig. 14). Based on the textured model, a complete set of orthophotographs has been generated at sufficient resolution to satisfy the requirements of many analyses (Fig. 13- left).

The main advantages of Terrestrial Laser Scanning over the other techniques of 3D data capture are: high speed, non-contact with the analysed object, both day or night data collection, direct acquisition of 3D coordinates with high accuracy and level of detail, and additional photorealistic color values (RGB) and intensity, which can contribute to a whole range of outputs, like a highly edited surface mesh, 2D and 3D drawings, orthophotos and rendered or virtual models, as indicated above.
However, TLS does not provide a unique solution to all recording tasks, to all types of elements of cultural heritage or, of course, at low cost or with an appropriate benefit-cost ratio. Therefore, in most cases, TLS is used with other data capture techniques, mainly digital photogrammetry and supported both in classical topography and GPS for georeferencing and control.

With regard to our experience in the survey of the Roman walls of Lugo with TLS as compared to other surveys of buildings and archaeological sites, obtaining the point cloud with such an amount of topographic data and such a high level of accuracy and resolution with respect to other techniques like classical topographical surveys or hand drawing has involved a significant reduction in fieldwork time and, therefore, in the cost of the budget. Yet, in many cases, TLS might be considered unnecessary and inadequate in terms of benefit-cost because of the size of the object and the purpose of the survey.

In this case, TLS is the best option because of the large size of the walls, about 2 km, and the need for a highly accurate and detailed geometrical documentation and digital data archive with which we can work in the future. As suggested above, the level of accuracy with the TLS is less than 2 cm. In addition, it is possible to combine different resolutions: very high-resolution measurement, 1x1 cm in some areas of complexity and interest, like wall gates, some areas of the parapet, towers and stairways, and lower-point density, 2x2 cm, for the rest of the wall.

With regard to the outputs obtained from point cloud processing, like the vector model, the orthophotos or the textured model, the post-processing of scanned data is particularly complex and achieving the required quality level for the results may take a long time. A line drawing to produce a CAD model is not an automatic process, insofar as the selection of vertices is a laborious and time-consuming task. Some efficient techniques have been reported in the literature for object recognition and automatic extractions. Yet, in single objects, further improvement is required to reduce the amount of time and labour in this process.

Finally, because the images needed for textured and virtual models obtained with TLS do not have sufficient geometric and radiometric quality, an external camera must be used. In addition, the field of view of the TLS, 60° in the vertical direction, both 38° up the horizontal and 22° down the horizontal, is an important factor that limits the capture of images. TLS is not as versatile as digital photogrammetry in terms of image data capturing and processing, but in this case terrestrial photogrammetry is not the solution, insofar as it is not possible to get the geometric requirements for the capture of images because of wall height. The height of these walls requires placing platforms or scaffolds for collecting the photographs, which is impossible around the walls. In this case, the solution is the use of rectified images obtained from the ground to the elevations and UAV images for the parapet.

Our experience in this project demonstrates that despite its benefits, TLS has also some drawbacks. Therefore, laser scanning is not a comprehensive and definitive solution for 3D modelling, at least in applications related to different modes of architectural and archaeological heritage. The potential of the combination and integration of terrestrial laser scanning with other techniques of 3D data capture, like UAV photogrammetry in this case, could be the solution for geometric documentation, particularly in cultural heritage applications, in which combining the best geometry of some techniques and the best radiometry of other techniques seems to produce the best results so far.

6. Conclusions

This paper describes the potential of the combination of terrestrial laser scanning and
UAV photogrammetry for the documentation of one of the most important monuments in the Spanish cultural heritage, the Roman Walls of Lugo. The integration of these two technologies provides massive and complete information for its record, analysis and visualization.

Nowadays, there is high demand for documentation of cultural heritage objects such as artefacts, sculptures or buildings. Terrestrial laser scanners are meaningful systems for deriving geometrical information, and UAV photogrammetry is a relatively new technology that has steadily emerged in cultural heritage. The lack of contact, the high accuracy and resolution of the 3D models obtained by combining these techniques and their ability to readily obtain measurements in inaccessible areas are some of their advantages. Despite the speed and accuracy of 3D measurements, data processing after capture is rather time-consuming and requires technical knowledge. Because both systems still require the support of other techniques such as GPS or conventional topography, none of them appears to be an integral solution.

According to the results obtained in this project, we argue for the synergy of both technologies, which allows for the use of the strengths inherent to both systems given the complexity of some heritage objects and the lack of a simple method that can provide a satisfactory solution under any measurement conditions. Assuming that the results of integration must be equal to or better than the results of the lack of integration, future research lines must focus on the improvement of the systems and tools used to manipulate and display the data obtained from the integration of both techniques.

References


Petti, F. M., M. Avanzini, M. Belvedere, M. De Gasperi,


Abstract:
The palimpsest nature of some rupestrial engraving is a drawback to the analysis of their motifs. The aging of engraving grooves makes them appear uniform, erasing diachronic information, and conserving only geometry. Graphic solutions based on light/shadow relationships are insufficient for relative chronology. A more efficient way to characterize “engraving groove families” is by distinguishing morphologies of strokes. We chose a set of engravings from El Mirón Cave as an optimal test case for 3D scanning. There is an accumulation of linear engravings on a block, fallen from the cave ceiling atop a Lower Magdalenian layer, which was covered by later Magdalenian deposits. Radiocarbon assays of levels pre- and post-dating them place the engravings around 16,000-13,000 BP. 3D data were collected with a structured light scanner with different fields of view at resolutions ranging from 50 µm to 280 µm. We present strategies based on semi-automated curvature extraction and other geometric issues.

Keywords:
3D Scanning, Rock Engravings, Upper Palaeolithic

1. The Engraved Block of El Mirón Cave

El Mirón Cave is located in Ramales de la Victoria - (ETRS89) 43° 14’ 42.76” N / 3° 27’ 9.16” W - in the upper valley of the Asón River in eastern Cantabria province, close to the border with Vizcaya. It is on the northern edge of the Cantabrian Cordillera, some 20 km from the present shore of the Bay of Biscay at an elevation of 260 m asl.

It is surrounded by peaks that reach and even exceed 1,000 m asl and is below los Tornos pass (920 m) that connects the Cantabrian Coast with the meseta of Old Castile in Burgos Province. The cave faces due west from a steep cliff on Mt. Pando and has a large mouth (20 m high x 16 m wide). The vestibule is 30 m deep x 8-10 m wide x 13 m high. This leads back to an inner cave that is accessible for another 100 m. The archaeological site is mainly located in the vestibule. It was discovered in 1903 but largely ignored since then, despite being adjacent to the well-known cave art sites of Covalanas and La Haza. Straus and González-Morales have been conducting systematic excavations since 1996, uncovering a long, rich cultural sequence extending from the late Middle Palaeolithic through the early Bronze Age, plus evidence of human visits in the Middle Ages, dated by 83 radiocarbon assays between 41,000 BP and 1,400 AD. The most important part of the sequence is the series of levels pertaining to the Magdalenián-Azilian cultural complex, Late Last Glacial. It is to this period that the engraved block relates.
One of the few large limestone blocks and by far the largest one found during the excavations is the object of the present study (Fig. 1, left), as its western face is decorated with engraved lines and its eastern face is stained with red ochre. The top of this block was visible at the start of the research, though largely buried in goat excrement and mixed sediments. Subsequent excavation revealed it to measure over 1.7 m long on a North-South axis about 1 m wide East-West and about 1 m thick. The northern end of the block had been undercut and possibly damaged by the clandestine digging of a ca. 25 cubic meter pit prior to 1996. Controlled excavation of the sediments immediately to the west of the block (squares V7, V8) revealed that the block had fallen atop Level 110 (as also shown by the southern face of the treasure hunters’ pit under the block overhang) and that it has subsequently been covered over by Levels 109-101. This means that the terminus post quem for the engravings on the block is the formation of Level 110 which is radiocarbon dated directly under or in the immediate vicinity of the base of the block to 16,520±40 and 16,130±250 (uncal.) BP, respectively, and corresponds to the end of the Initial Cantabrian Magdalenian.

The block had fallen from the cave ceiling at the angle it forms with the rear wall of the vestibule. Its weathered "outer" face (the former ceiling) landed atop Level 110 and the flat inner face that had sheared off along a plane of weakness in the bedrock landed in a position tilted at an angle oriented toward the cave mouth, such that sunlight reaches it at the end of the afternoon in summertime. It was this flat face that was engraved and subsequently covered over by Lower, Middle and Upper Magdalenian layers dated between about 15,000-12,000 (uncal.) BP, and finally by sediments of Holocene age. The Middle and Upper Magdalenian levels provide the terminus ante quem dating for the engravings. The eastern face of the block was painted red, apparently in relationship to the secondary burial of ca.100 ochre-stained bones of a young adult human in sediments also impregnated with red ochre rich in hematite crystals during the Cantabrian Lower Magdalenian around 15,600 (uncal.) BP (Straus and González-Morales 2011).

The engravings on the smooth, west-facing side of the block are all linear and do not seem to be representative of anything recognizable at the present time. There are two generations of lines: shallow + thin and deep + wide. By extrapolating the slope of the relevant archaeological levels, it has been deduced that similar linear engravings, as well as a fine engraving of a horse on the rear wall of the vestibule close to the engraved block were probably also executed during the Lower-Middle Magdalenian (García Díez et al. 2012). Although non-figurative, the engravings on the block are among the most precisely dated yet known for the European Upper Palaeolithic, rivaled only by such sites as Le Placard in west-central France and Ambrosio in southern Spain. The Lower Magdalenian levels in El Mirón (as in other sites of the region) are extremely rich in lithic and bone artefacts, remains of red deer, ibex and salmon, hearths with fire-cracked rocks, personal ornaments such as perforated shells and teeth, and works of portable art, notably red deer scapulae engraved with images of red deer and other ungulates, typical of this period in the central Cantabrian region (González-Morales and Straus 2009).

This paper is organized as follows. First, we proceeded with an archaeological contextualization of the engraved block of El Mirón. Second, we address the issue of the insufficiency of graphic solutions based on light/shadow relationships for resolving the problem of palimpsests of engravings. Then, we describe our approach, based on the 3d scanning of the grooves, and subsequent morphological feature extraction, analysis, and comparison. As a work in process, preliminary results are presented to demonstrate the potentialities of the used methods, tools, and techniques. Finally, a conclusion is given.
2. The Question of Rock Engravings

The difficulty in reading and interpreting surfaces with high densities of superimposed engravings – so-called palimpsests (Ripoll 1972) – is that we only have their geometric traces, the form of the lines (Mélard 2010). What gave sense to the act of repeatedly engraving the same space, overlapping marks and, apparently, the annulling of previously drawn motifs, was the fact that the grooves also displayed colour as they were originally made. The trace of each design cut through the surficial patina of the rock being engraved, acting like a coloured pencil. With the passage of time, the grooves would take on a patina like the rest of the rock and they would become camouflaged. In this sense, the surface once again would become attractive and practical for the engraving of new designs. It is also possible to think that the surface was consciously smeared to wipe out the contrast of different patinas.

Until relatively recently, before the impact of computer applications in archaeology, the method used to document engravings on-site was to use low-angle lighting, directing several beams at as many grooves as possible. If it was the goal of this technique to capture all the images by photography, it was often impossible to do so with a single shot. The strategy was to set up many spotlights, seeking a balance between light and shadow to make visible as many of the designs as possible. But the more complicated the lighting system, the less natural the results became; engraved panels were never seen that way in real life. What we, the archaeologists, lacked was an important part of the graphic information, colour.

Like any other research methodology, computer applications do not necessary make the results any more “objective”, but they help us to quantify the subjective choices we make as researchers. Nonetheless, in the case of the analysis of prehistoric art – or “prehistoric graphic expression”, to avoid the controversial term “art” –, computer application tools have frequently been incorporated in the documentation process in a rather peculiar and, doubtless, unstructured way. Especially in the last decade, common consumer photo manipulation applications are the order of the day. It is not necessary to name the computer programs for publishing and retouching photographs with which most of us compose our graphic material. The most critical aspect of many of these computer solutions is the difficulty for other researchers to replicate any process done by someone else, or even for oneself to repeat one’s own exact work. Generally, documentation of rock art studies do not specify the process followed or, at least, it is very difficult to find a coherent description of the variables and choices made to achieve the final graphic product (Rogerio-Candelera 2008). Especially on paintings, the researcher seldom pursues a systematic workflow pattern to build the graphical process. By contrast, it is based on a chain of quasi chaotic trial and error method until the “emerged” image satisfies, i.e., satisfies his visual experience, obviously subjective.

On the other hand, the incorporation of these new tools has been done only to facilitate the operations of image capture as they had been conceived of before the digital age, without any reflection about the nature and totally new possibilities of the new tools. We continue to make “tracings” by hand, but using a computer mouse rather than a pencil (Collado 2009). With reference to rock engravings, in the best of cases one applied renderings of solar studies or the play of lights and shadows within a geometric digital model (Cassen and Robin 2010). This doesn’t mean that it is an inadequate technique. On the contrary, the recreation of the visual experience by means of virtual maps of lights and shadows is a natural means for the understanding of engraved designs. Nevertheless, as mentioned above, this strategy is insufficient for resolving the problem of palimpsests of engravings. However, aided by the vast calculation capacities of computers,
we can organize new modes of documentation oriented toward the characterization of engraved grooves based on their specific geometry. Of course, we cannot recuperate the original colour, but can help obtain hierarchical images of engraving styles, forms of working, evidence of manual habits. This was our goal in doing the experiment with the Magdalenian engravings on the block in El Mirón Cave.

3. Methods

In this context, our approach towards the study of the grooves of the engraved block of El Mirón comprises the following steps. First, to capture the surface’s geometric irregularities of the entire block by using a 3D scanner with different fields of view (FOV), in order to assess their overall performance, and to test their range of applicability, i.e., the required level of detail to be used in other similar engravings. For their study it is essential to measure, to compare and to classify the various attributes of the forms of archaeological materials, as much as to quantify them, because these allow to describe its (ir)regularity and to some extent making possible the study of its causes (Barceló 2010). Then, to post-process the captured data and generate a 3D digital surface model. Next, to proceed with a semi-automatic detection of the surface irregularities, and analysis of its curvatures. After that, to isolate the strokes and extract several profiles. Ultimately, we aim to distinguish different geometries of strokes – namely, 3D areal and 2D profile parameters, morphology, depth, length, width – and analyze statistically any potential stroke clusters, in order to attempt to isolate the motifs and suggest possible types of engraving tools.

3.1 3D Scanning and Morphological Data Extraction

The entire surface with the engravings was scanned with a 3D structured light scanner (SmartSCAN 3D Duo System, Breukmann), with three different stereo FOV: the 90 mm, the 150 mm, and the 450 mm.

There are several technical reasons that can explain the complexity of scanning, namely the relative location, distribution, and type of engravings; environment lightning conditions; and hardware-software issues. To minimize measurement error, a sufficient amount of overlapping area between scans is needed. To overcome most of the self-occlusion problems, we used multiple viewing angles during scans, by changing the scanner’s position in respect to the engraved surface of the block. Regarding to the scanning environment lightning conditions, for better results, capturing data with structured light scanners should be done in a dark environment – when outdoors, preferably during the night – because the more contrasted the projected and reflected patterns are, the more consistent data we get. Likewise, it is important to realise that either missing data – which leads to holes in the mesh –, noise data – which leads to erroneous data, causing lack of

Figure 1. The engraved block of El Mirón’s Cave (left). Preparation of the 3D scanning workspace (right).
data consistency in some areas – filling holes, filtering and certain parameter’s settings during scanning or post-processing may conceal or distort relevant data, such as engraving grooves (Moitinho de Almeida and Barceló 2012). Inasmuch we were only concerned with the geometric data of the grooves, we decided not to capture any image texture during scanning.

The 90 mm set of lenses – which, according to the manufacturer, has a resolution of 50 µm – required a total number of 77 scans, in 3 parts later merged. Throughout this process, in view of the fact that the provided fabric curtains (Fig. 1, right) didn’t entirely prevent the entrance of indirect sunlight, and the intensity of light entering the cave varied dramatically during the day – sunny to cloudy and rainy day –, the parameters of the scans had to be constantly readjusted and frequently repeated, inevitably causing some noise data. This problem was only overcome before scanning with the 150 mm and the 450 mm sets of lenses, by darkening the scanning area homogeneously with more curtains. Unfortunately, due to the short amount of time available, we were not able to repeat the shortest FOV scans.

An important factor to take into account is that the higher the FOV, the lower the number of scans needed to cover the entire area. However, the resolution also tends to diminish leading to less fine density 3D meshes. Therefore, as to the 150 mm FOV, it required 26 scans at a resolution of 90 µm, whereas the 450 FOV required only 18 scans at a resolution of 280 µm. Both capture and post-processing of the three dimensional digital data – i.e., from point cloud to 3D digital surface model – was done using the scanner’s software Optocat 2009R1 v.8.00.28-1520, x64 Bit Edition.

In this preliminary study we decided to focus on a small area of the engraved block, because of the large concentration and apparent variety of grooves herein (Fig. 2), and in one of its strokes (Fig. 3).
Analyzing the amount of curvature at a surface point allows identifying edges and patterns, in other words the bending energy measure of surface irregularities and eventual engraving grooves. Curvedness is specified by a quantitative measure, and its value is invariant to translation and rotation. For this we used RapidForm XO Scan 2010 (INUS Technology) software, which permits to proceed with their extraction and analysis in a semi-automatic way. Since each FOV determines the resolution of the captured geometries of the micro topography, we had to change the parameters of the maximum curvature until we got best results for an efficient analysis (Fig. 4). The colourmap is from red (convex) to blue (concave), corresponding to high to low values of the curvedness, whereas grey colour indicates planar patches.

After selecting the engraved groove of our study area, we isolated it to thereafter be able to extract the profiles for, in this case, five parallel equal cross-sections – perpendicular to a plane tangent to the upper surface, and starting from the same aligned position –, by means of automatic multi-slice technique (Fig. 5). The sampled profiles of each FOV were then overlapped and positioned (Fig. 6), in order to be able to proceed with further analysis of the morphological data – including calculating the mean profile –, as well as typological classification and quantitative comparisons (Karasik and Smilansky 2008).

4. Results and Discussion

Even though the 90 mm FOV scans have a considerable amount of data noise – when compared with both 150 mm and 450 mm FOV – it seems clear for us that this set of lenses permits capture of engraved stroke morphology...
The 450 mm set of lenses proved to be not suitable at all for our specific study, when compared with both 90 mm and 150 mm sets of lenses. On one hand, it is obvious the inefficiency in capturing the overall irregularities of the surface — some of the strokes weren’t even captured. On the other hand, the inaccuracy of a groove’s morphology is evident, as it becomes even wider, smoother, shallower and less sharp, in a way that it becomes almost impossible to distinguish one stroke from another, and thus proceed with further geometric comparisons.

As this is a work in progress, presently we are continuing to isolate the strokes and, at the same time, extracting the profiles for equal sections of each of the engraved groove of our study area. Later, we aim to extend this to the entire engraved block. As part of the procedure, we will subsequently overlap these profiles and analyze and characterize the geometric data. The next step will consist in statistically try to group them into stroke clusters. In the end, we wish to distinguish different types of strokes in order to attempt to isolate the motifs and determine possible types of engraving tools.

**Aknowledgements**

Excavations has been authorized (and partially financed) by the Regional Government of Cantabria, and funded by U.S. National Science Foundation, Fundación M. Botín, L.S.B.Leakey Foundation, Ministerio de Educación y Ciencia, National Geographic Society, University of New Mexico, Stone Age Research Fund (Jean and Ray Auel, principal donors), and with material support by the Universidad de Cantabria and the Town of Ramales de la Victoria.

The authors wish to thank the Institució Milá i Fontanals of the Consejo Superior de Investigacion Científica (IMF-CSIC), for providing the 3D scanner, without which it would not have been possible to capture the 3D digital models of the engraved block. This

![Figure 6](image_url)
research also benefits from Vera Moitinho’s Ph. D. grant from the Fundação para a Ciência e Tecnologia (FCT), Portugal, co-funded by the European Social Fund.

References:


Meshlab as a Complete Open Tool for the Integration of Photos and Colour with High-Resolution 3D Geometry Data

Marco Callieri, Guido Ranzuglia, Matteo Dellepiane, Paolo Cignoni and Roberto Scopigno
Visual computing Lab, ISTI - CNR, Italy

Abstract:
Even if a precise, detailed geometry is essential for the study and documentation of a Cultural Heritage (CH) artifact, lots of information also comes from the colour and appearance data. Since a complete reconstruction of the optical properties of an artifact is still hard (or impossible) to achieve, what is normally done is to rely on colour data coming from photographic dataset. In this paper, we will describe the tools that MeshLab offers for the generation and processing of colour information on a high-resolution 3D model. We will show the basic operations available and present actual projects where these tools have been successfully used. The availability of such advanced colour management tools in an Open Source tool is important for the archaeological and CH community. MeshLab is a complete, state-of-the-art open source software for the creation and manipulation of high-resolution 3D models starting from real-world 3D scanned data, at par with the specific commercial tools available on the market.

Keywords:
MeshLab, 3D Scanning, Color Mapping, Photographic Registration, Texture Mapping

1. Introduction

In digital 3D models, the reproduction of apparent colour, obtained by mapping photographic information on a 3D model is still the most common approach to describe the visual appearance of an object, especially in the cultural heritage field.

While a more complex representation of the material of the object, like a Bi-Directional Reflectance Distribution Function (BRDF) would have been much more effective, it is often impractical (due to time and resources constraints) to employ the sampling techniques needed to acquire such complex representations.

The advance of 3D scanning hardware has been, in the last few years, quite significant; on the market are nowadays available many different devices, able to capture objects with a wide range of materials/sizes/complexity. Unfortunately, the same technological advance had not much effect on the ability of devices to capture the colour of scanned objects. Most 3D scanners do not capture colour. In the devices with such sensors, the generated colour is generally poor in terms of resolution (the colour sensor is often an embedded camera derived from phone/mobile photographic devices) and photometric quality (no explicit calibration is possible, shooting parameters are chosen by the device and no manual mode is available, the software does not support saving in uncompressed formats, hidden colour processing is done in the driver software). It may be enough as a preview or for some industrial processes, but not for the needs of Cultural Heritage operators.

For this reason, a 3D Scanning campaign is usually supported by a photographic campaign, since the quality of the colour acquired by scanner is not always sufficient, or the acquisition is performed in light conditions that are bad for photography (but do not influence the scanner).
The task of effectively using photographs to generate documentation for 3D models, and to generate a colour mapping of the 3D surface is still one of the most problematic steps in the processing pipeline of 3D scanning data.

2. Previous Work

While most of the intermediate steps in the colour mapping pipeline have been active topics of research, it is difficult to find examples of complete software tools able to manage in a complete way the process of applying colour to 3D models. Each research group active in this field has produced software prototypes for their methods; these tools, however, are generally available by request, and it may prove to be difficult to use them to build a complete software pipeline.

Some commercial 3D tools, in particular those oriented to the management of data produced by 3D scanners, like RapidForm or Geomagic, do implement (in part or completely) the colour mapping pipeline. MeshLab, on the other hand, does implement the entire colour management pipeline.

The first step in the colour projection pipeline is image registration, since in most cases, the camera parameters associated with each image are not known in advance. Several automatic (Ikeuchi et al. 2007; Brunie et al. 1992; Lensch et al. 2000; Wolberg and Zokai 2006) and semi-automatic (Franken et al. 2005) methods for image-to-geometry registration have been proposed. They are mainly based on an analysis of the geometric features of the model (e.g., silhouette and orthogonality), or on some input given by the user (2D-to-3D correspondences).

The most correct way to represent the material properties of an object is to describe them through a reflection function (i.e. BRDF), which attempts to model the observed scattering behavior of a class of real surfaces. A detailed presentation of its theory and applications can be found in (Dorsey et al. 2007). Unfortunately, state-of-the-art BRDF calculation approaches rely on controlled and complex illumination setups (Lensch et al. 2001; Debevec et al. 2000): this limits their application in the context of complex scanning projects (big artifacts and on-the-field acquisitions, such as those performed in museums).

A less accurate, but more robust, solution is the direct use of images to as a source of colour information for 3D surfaces. In these cases, the apparent colour value, as sampled in digital photographs, is mapped on the digital object’s surface by registering these photographs onto the 3D model (by estimating the camera parameters), and then applying an inverse projection. In addition to other important issues (cited in sections 5 and 6), such as the image registration and how to store colour information, there are numerous difficulties in selecting the correct colour when multiple candidates are in different images. Some of these are: how to deal with discontinuities caused by colour differences between photos that cover adjacent areas, and how to reduce illumination-related artifacts, e.g., shadows, highlights, and peculiar BRDF effects.

To solve these problems, one group of methods selects, for each part of the surface, a portion of a representative image following a specific criterion - in most cases, the orthogonality between the surface and the view direction (Callieri et al. 2002; Bannai et al. 2004). However, artifacts caused by the lack of consistency between overlapping images are visible on the borders between surface areas that receive colour from different images. These can be partially removed by working on the border between two images (Callieri et al. 2002; Bannai et al. 2004).

Another group of methods “blends” the contribution of all the images by assigning a weight to each one or to each input pixel (this value expresses the “quality” of its contribution),
and selecting the final surface colour as the weighted average of the input data, as in (Pulli et al. 1998). The weight is usually a combination of various quality metrics (Bernardini et al. 2001; Rankov 2005). In particular, (Callieri et al. 2008) presented a flexible weighting system that can be extended in order to accommodate additional metrics.

Another important step in the colour management is the creation of a parametrization for the 3D model; this is still an active research topic, and a large number of papers on surface parameterization have been recently published. A good survey of this field can be found in (Sheffer and Praun 2006).

3. The Tool

MeshLab (Cignoni et al. 2008) is an open source tool for the visualization and processing of 3D models. It is oriented to the management and processing of large, unstructured triangular meshes and pointclouds, and it provides a set of tools for measuring, checking, cleaning, healing, inspecting, rendering and converting 3D meshes. MeshLab is freely available, distributed under the GPL licensing scheme and it is available for all the major platforms (Windows, MacOS, Linux).

Born as a university project, MeshLab has steadily grown in features and usability, reaching more than 200,000 downloads in this year. MeshLab is used by hundreds of research groups and industries, and by thousands of 3D hobbyists. Featuring various state-of-the-art 3D processing algorithms (often implemented by their academic authors), it represents a solid and free alternative to commercial tools for 3D scanning data management.

Processing 3D scanned data is not a simple task; most scanners do have a bundle software able to manage the data acquisition and manage simple to medium-sized projects. These tools, however, only work with a specific device. Processing of large and complex digitization projects, or mixing data coming from different devices does, in general, require specific commercial software tools, available on the market. MeshLab does implement fully the 3D scanning and colour pipeline, is able to process a large amount of data, does accept 3D data coming from different devices and it is readily available to users.

The design of MeshLab is based on a modular structure: around some core structures, able to manage high resolution 3D data and provide basic rendering functions, all the functionalities of the tool are implemented as individual operations, each one independent from the others.

MeshLab relies on a wide variety of mesh processing functionalities, which are exposed as a large set of self-contained filters that take as an input one or more meshes and some user specified parameters. MeshLab offers around two hundred different filtering operations, categorized into menus according to a few keywords (remeshing, cleaning, sampling, texture, colour processing, quad mesh processing, point clouds, etc). The simple functional structure of filters is implemented through a plug-in framework that hides all the GUI coding allowing for a very simple development of additional plug-ins by new developers.

This “Swiss Army Knife” approach is also one of the most powerful features of MeshLab. Providing many different self-contained instruments, in many cases with multiple alternative filters that do the same operation, makes possible for the user to build a custom sequence of editing operations to solve specific problems of specific datasets.

While having a single-push button tool may be easier for beginners, a fully configurable and flexible one is absolutely necessary for people working on very diverse kind of objects, where each scanning project is different from
the previous one, as often happens in the cultural heritage field.

4. Raster Layers

In order to support the management of photographic information, a new data type has been added to the MeshLab: the raster layers. Mimicking what is done for 3D meshes (contained in the geometry layers), the system now supports the loading of a series of images.

These layers do contain a raster image, plus the camera parameters needed to establish a correspondence between the 3D geometry and the 2D image. These parameters describe the position and orientation (extrinsic parameters) and internals of the camera, like sensor size, lens distortion and focal length (intrinsic parameters) at the moment of the shot. By obtaining these parameters it is possible to reconstruct the perspective projection that created the photo. This opens up two possibilities: being able to see the 3D scene through the same camera that took the shot (thus, exploring the photographic dataset spatially), and project back the colour information onto the 3D model (to generate colour mapping).

5. Photographic Alignment

If the photos have not been acquired by the 3D scanner (thus, being automatically registered with the 3D information), some processing steps are needed to link them to the geometry, in order to obtain a set of calibrated images, and to eventually project the colour information on the geometry.

This first step is the alignment of the images on the 3D model: this is obtained by estimating the camera parameters associated to each image. This stage may also be called Image Registration or Camera Calibration and Orientation.

This is not a simple task, since the estimation of the parameters, which is a geometrical and mathematical task, is a badly conditioned problem. Some of the parameters are mathematically linked, and it is easy to obtain a camera calibration which is not completely correct, ending in a local minima.

Since there is no simple way to check the correctness of the alignment, a visual feedback must be given to the user in order to assess the quality of the recovered data.

The implementation of photographic alignment inside MeshLab has been designed in order to cope with these problems.

The most robust approaches to image alignment are based on the setting of correspondences between the 3D model and the image (Franken et al. 2005). Yet, this point picking operation may be difficult, depending on the characteristics of the 3D object. Hence, a different approach, based on Mutual Information (Corsini et al. 2009) was applied. This is based on the calculation of a statistical measure of correlation between the image and a rendering of the model.

The Mutual Information approach has been implemented as an additional feature of MeshLab. Using an easy to use visualization, it possible to have a direct feedback about the quality of the alignment. This alignment strategy works on triangulated 3D models but also on point clouds, making it quite versatile.
Another alignment method, able to combine user-picked reference points and the Mutual Information is being added, and will be released in a future version of the tool. This addition will make the alignment of difficult datasets easier, at the price of a more complex user interaction.

Photographs aligned in this way may be spatially-explored directly in the 3D space, instead of browsing a folder on the disk, looking at the photos like see-through transparencies suspended in space or projected one by one onto the 3D surface.

An interesting possibility offered by this strategy is to use unconventional images, like photos with annotations, historical photos, hand drawings or sketches or even near-visible lighting photos (ultraviolet, multispectral, infrared, thermography, etc.). This possibility of spatial exploration of the geo-referenced photographic set is a powerful tool to effectively browse a collection, and has multiple uses in the CH domain.

However, in many applications, it is necessary to have colour information mapped onto geometry. With this aim, MeshLab offers different colour mapping tools to better cope with the different needs of the various datasets. By using the colour data from the calibrated images, it is possible to generate detailed, artifact free ‘per-vertex’ colour encoding, to fill an existing texture parametrization or to generate entirely new texture mapping, driven by the photographic coverage.

6. Colour Storage

The first problem in colour mapping is where to store the colour information on the 3D object.

While the management of 3D geometry is more or less standard among the various software for the 3D creation, editing and rendering, the support of colour information is still quite different from one tool to another. In addition, different file formats support different methods of colour storage, adding complexity to the task.

The two main methods of colour storage are per-vertex (each vertex of the mesh has an associated RGB value) or texture mapping (3d object has a parametrization, which associate a pixel in a 2 image to each point on the 3D surface, and the colour is stored in an image, called texture map).

Per-vertex encoding works well for highly dense models, such as the ones produced by 3D scanning; this because the colour detail is limited by the resolution of the geometry. It is a compact and effective way to store colour, it is simple to use and works with most 3D modelling and rendering tools.

On the other hand, texture map is a more standard solution, it decouples the resolution of the geometry from the resolution of the colour detail and it is supported (with small variations) by all 3D software.

While the texture approach seems more usable and standard in the field of 3D graphics, it requires the 3D geometry to have a parametrization (a correspondence between a point on the 3D surface and a point on the 2D plane). This requirement is difficult to fulfill in the case of digital models coming from...
Meshlab as a Complete Open Tool for the Integration of Photos and Colour with High-Resolution 3D Geometry Data
Marco Callieri et al.

3D scanning. These models, because of their geometric complexity and their unstructured nature (they are, basically, just a set of unorganized triangles, often with topological problems), are quite difficult to parametrize, imposing serious limitations on the texture mapping approach, especially for complex 3D models (more than 1 million triangles).

To solve this issue, MeshLab implements a series of algorithms able to generate a texture parametrization for the input 3D models. MeshLab can generate a quite trivial parametrization (each triangle packed on its own in the texture space), usable on simple objects, or an extremely advanced global iso-parametrization, with optimal properties in terms of low distortion and continuity of mapping (see (Pietroni et al. 2010) for more details).

A third parametrization method relies on the use of calibrated images to map a 3D surface to a 2D plane. It works by subdividing the mesh in parts, each one that projects correctly one of the input photos: each of these subparts is then flattened on the texture space by applying the projection associated to the corresponding photographic camera (more details in Section 8).

Finally, a general-purpose, automatic mesh flattening tool has been recently added, able to work on arbitrary 3D models, producing surface parametrizations that are well suited for texture mapping.

Again, the strategy has been to include multiple operations to compute the same kind of data, in order to cope with the specific characteristics of the projects the user has to work with. Every one of the parametrization tools has strengths and weakness, and the user may choose the one that better suits their needs, without affecting the other stages of the workflow.

7. Weighted Blending

Given a 3D surface and a set of calibrated images, it is possible to project colour information from the images to the surface. However, the hardest problem is that for each point on the surface there are many photos that are possible contributors. Due to lighting inconsistencies, unstable camera settings and sensor noise, the colour from the different sources is never the same, and it is not easy to choose which colour map onto the surface.

While a simple blending would blur the final colour; a better solution is to intelligently blend all contributions according to their quality. This will ensure consistency of colour and preserve the detail of information present in the input photos.

The MeshLab filter implements an improved version of the weighted blending method described in (Callieri et al. 2008), with a more efficient and fast computation, a smaller memory footprint and some tweaking in the

Figure 3. An example of raw geometry, and of the increased realism obtained through photographic colour mapping.
weighting functions.

The method can calculate a colour for any point over the surface, by estimating the most correct colour as a weighted mean, which takes into account various quality metrics. Each pixel of each of the input images has a quality value associated to it: this quality value is calculated using multiple metrics, like the distance of the camera to the sampled points (closer is better), the viewing angle (when the camera is more orthogonal to the surface, it is better) and the proximity of the pixel to critical points on the photo (border pixels are bad, and so are pixels close to depth discontinuities). This meaningful and smooth weighting system ensures that weighted blending produces continuous and detailed colour mapping.

This colour mapping technique does work to both generate per-vertex colour (by evaluating the colour blending for each vertex) or to fill an existing parametrization (evaluating the colour blending for each texel). This makes the method flexible, enabling the user to choose the most appropriate output format. This method produces coherent colour, since all redundancy between images is effectively used. However, since blending is involved, a small loss in higher frequency colour details is unavoidable. The method scales well with the size of dataset, both in terms of number of input images, and in terms of the complexity of the 3D model to be mapped.

For these reasons, this method is better suited for datasets composed of a large number of images, and/or complex datasets. In our experience, we were able to map 100+ images over 3D models of 10-20 millions triangles.

8. Photo Stitching

An alternative method to map colour on a triangulated model is to build a texture map, stitching together parts of input photos.

Using the geometry and the aligned photos, it is possible to find the photograph that gives the best colour information for each part of the model. A texture map can be assembled by stitching together sections of the photos, according to this partition.

To this aim, MeshLab does implement the texturing method described in (Callieri et al. 2002), adding various optimizations and new features, like a better quality evaluation when assigning the photos to each part of the mesh, a more stable end efficient texture packing, and a more effective strategy to uniform colour.

The first step is to create a parametrization for the 3D model. The model is analyzed in order to determine which photograph best captures each triangle of the 3D surface. The “best” photograph for each part of the model is determined, again, considering the metrics described for the colour blending process (distance, angle and position in image space). In this way, the model is divided in various patches; pieces of the surface that are well represented in one of the aligned photos.

Then, each patch is projected on the corresponding image (using the camera parameters recovered during photo alignment), and the chosen part of the photo is cut and pasted onto the final texture. A final step aims at smoothing out the lighting inconsistencies (which are always present, across sets of photographs), by comparing the borders.

Figure 4. Colour mapped using weighted blending with per-vertex encoding.
between patches, and propagating this lighting correction across the image.

The same filter can also be used simply to generate the photo-driven parametrization of a 3D model, which can then later be filled using the weighted blending or by transferring colour information from a per-vertex encoding or another texture parametrization. This method has the advantage of producing sharp colour mapping (since there is no blending between the various photos). However, the method does not scale well with the size of the input dataset: too many photos produce a very fragmented texture, bad for storage, compression and probably with artifacts on borders; a too complex 3D model (more than 2 million triangles) may cause the algorithm to fail. Considering all these facts, this method is more suited to situations in which there are few, good pictures, and the size of the model is not too large.

9. Colour Editing

MeshLab implements a series of filters to apply colour correction to the whole model, these corrections are similar to the ones found in many image editing tools, and are quite useful to fine-tune the mapped colour. Using simple parameters it is possible to edit the gamma, saturation, contrast, levels, white balance, hue and brightness.

As mentioned above, colour management does change a lot from one software to another. To be able to exchange data between different software tools, conversion is often required from per-vertex colour encoding to texture mapping, texture re-parametrization, and colour data transfer across different models and textures.

MeshLab does provide different filters which can be used to obtain such functionality; by exploiting these operations it is possible to build an effective pipeline of software tools.

Finally, MeshLab includes and implementation of a “painting” filter. This can be used to apply colour to the 3D surface using a paint-like interface. This filter implements basic paintbrush settings, with transparency and hardness, colour noise, clone tool, eyedropper and colour smoothing. It is simple to use for anyone with a little experience with photographic editing tools.

The interface is quite simple, when compared to more complete tools like Z-Brush or DeepPaint 4D, but it is able to work on very large unstructured 3D meshes of many millions of triangles, which is definitely helpful in the CH domain.

This paint feature may be useful in many different situations, like correcting small problems of colour mapping (remove aliasing/ghosting, correct shadows of highlights, clone out errors in photos projections), annotating a 3D mesh (paint areas, drawing on the 3D surface), or even in helping art historians to reconstruct possible original colour of a damaged artifact.

The reconstruction of a faithful representation of the actual colour of an artifact is often just a first step towards the understanding of the original appearance of the object. Almost always, the current state of an archaeological find is degraded: to propose possible reconstruction of the original colour of
the object is a normal activity for archaeologists and art historians. This task, however, is generally carried out using photographs or drawings of the artifact, while it would be possible to do it directly on the 3D surface.

The Surface Painting tools provided by MeshLab may be used to colour 3D surfaces using paint-like interface. By using this editing tool, it was possible to produce such proposed colour reconstructions, directly on high-resolution 3D geometries.

10. Conclusions

We have described here the new functionalities of the MeshLab platform for the management of colour information for high-resolution 3D models.

MeshLab is a complete open tool for the management of colour information, when working on 3D scanned digital models. Being an open source tool, it is an extremely valuable instrument for the CH field, not because it is free, but also because it gives the user full access to the implementation details (ensuring a good knowledge of data provenance), and supports the use of open data formats.

Our belief is that, in order to offer a usable tool to users, it is important not only to provide a single, integrated tool with all necessary functionalities, but also to make the various functionalities as configurable and customizable as possible. The availability of multiple alternative ways to perform the same step ensures that the user will be able to choose the best solution to cope with the specific characteristics of their projects.

Plans for the future development of the tool will follow the same model of providing “multi-purpose” tool that has been at the core of the philosophy of the project. We will include in the tool other alternative functions to cope with the various steps of the colour processing; image alignment, parametrization and colour mapping. We aim to offer the user a powerful arsenal of tools, sufficient to cover the diverse situations that arise in the processing of 3D models for the CH field.

We will then devote some effort to minimizing the colour mapping artifacts due to small misalignments, and to make the tool even more scalable with respect to large datasets.
References


Enhancing Surface Features with the Radiance Scaling Meshlab Plugin

Xavier Granier
Inria Bordeaux Sud-Ouest and Inria LP2N, France

Romain Vergne
University of Giessen, Germany

Romain Pacanowski and Pascal Barla
Inria Bordeaux Sud-Ouest and Inria LP2N, France

Patrick Reuter
Inria Bordeaux Sud-Ouest, University of Bordeaux Segalen and Inria LP2N, France

Abstract:
We present the new Radiance Scaling plugin for Meshlab. This rendering technique allows depicting shape through shading via the modification of light intensities around specific features. The major idea is to exploit the fact that light intensities and surface curvatures are correlated. Hence, our technique modifies the shading according to surface feature variations in order to enhance shape details like concavities or convexities. As it works in real-time on modern graphics hardware, surface features can be inspected interactively. Recently, Radiance Scaling has received major interest in the French Archaeology research community, in particular for enhancing details in carved stones and thus improving their legibility.

Keywords:
Meshlab, 3D Shape Depiction, 3D Surface Analysis, Feature Highlighting

1. Introduction

The 3D acquisition of cultural heritage artifacts is becoming more and more common. Most often, the surface of the artifacts is either directly acquired by 3D laser scanning, or obtained by 3D photogrammetric reconstruction, followed by a 3D point to mesh conversion. One of the first motivations for 3D acquisition was to digitally preserve cultural heritage. More recently, it has been shown that for broken artifacts, the separate 3D acquisition of each fragment may offer the possibility to virtually reconstitute the past by computer aided reassembly.

During the virtual inspection of the acquired artifacts, the maybe most important element is the 3D visualization. Nowadays, the advances of computer hardware and especially the graphics cards make it possible to render photo-realistically the meshes of 3D models consisting of millions of polygons.

We are convinced that the inspection of 3D models is further improved when they are visualized with modern 3D expressive visualization techniques. With the exaggeration of several surface features according to the 3D model geometry, these techniques may better depict shape characteristics. Emphasizing meaningful portions of the surface and hiding the meaningless ones creates a far more legible pictorial representation, and removes possible ambiguities.

In this work, we present a specific expressive visualization technique called Radiance Scaling (Vergne 2010) that we integrated as a plugin for the 3D model processing software Meshlab. This rendering technique allows depicting shape through shading via the modification of light intensities around specific features like concavities and convexities. The theoretical foundations of Radiance Scaling introduced by Ramamoorthi et al. (2007) demonstrate...
through first-order analysis that curvature and light intensities are correlated. Radiance Scaling works in real-time on modern graphics hardware, making it feasible for an interactive inspection in Meshlab.

We believe that the choice to integrate Radiance Scaling as a plugin in Meshlab is the perfect means for a widespread use, because we observed a growing interest in Meshlab, especially in the digital cultural heritage community.

This paper is organized as follows. In Section 2, after introducing Meshlab, we show some related work concerning expressive visualization with a special interest for cultural heritage applications. Then, in Section 3, we present the Radiance Scaling plugin with all its options. In Section 4, we discuss a case study on a concrete example of using the Radiance Scaling plugin for deciphering inscriptions on an epitaph found near St. Emilion, France. Finally, we conclude with directions for future work in Section 5.

2. Related Work

2.1 Meshlab

Meshlab is an advanced 3D model processing software system for the processing, editing, cleaning, healing, inspection, and the visualization of 3D models.

It is particularly interesting for the research community for several reasons. From a general point of view, it is distributed freely, and as open source, and it is portable to a number of operating systems, and deals with a large variety of 3D mesh formats (e.g., PLY, STL, OFF, OBJ, 3DS, VRML 2.0, U3D, X3D and COLLADA), and even point clouds (e.g., imported from photogrammetric reconstruction such as Photosynth).

Meshlab is an extensible system, and from a computer scientist point of view, any developer can contribute to its improvement. The open architecture makes it possible to write I/O plugins for new 3D formats, to develop editing plugins for advanced mesh processing, and to propose new rendering styles, even for programmable graphics hardware. We made use of this latter for the integration of our new radiance scaling technique.

Meshlab has been developed since 2005, and the user and developer community is growing steadily, and especially in the digital cultural heritage context, where it has already proven to be beneficial.

2.2 Expressive Visualization

The depiction of object shape has been a subject of increased interest in the Computer Graphics community since the work of Saito and Takahashi (Saito 1990). Inspired by their pioneering approach, many rendering techniques have focused on finding an appropriate set of lines to depict object shape. In contrast to line-based approaches, other techniques depict object shape through shading. Maybe the most widely used of these is Ambient Occlusion (Pharr and Green 2004), which measures the occlusion of nearby geometry. Both types of techniques make drastic choices for the type of material, illumination and style used to depict an object: line-based approaches often ignore material and illumination and depict mainly sharp surface features, whereas occlusion-based techniques specifically depend on sky-like lighting environment.

Among the recent techniques, the work of Vergne et al. (Vergne 2009) is receiving raising interest thanks to its simple integration into a large range of shading techniques. For Radiance Scaling (Vergne 2010), the key observation is that reflected lighting variations are correlated to surface feature variations. For example, consider a highlight reflected from a glossy object: by increasing reflected light intensity in convex regions and decreasing it in
concave ones, the highlight looks as if it were attracted towards convexities and repelled from concavities. Such an adjustment improves the distinction between concave and convex surface features, and it does not only take surface features into account, but also material characteristics.

3. The Meshlab Plugin Radiance Scaling

For Meshlab (Fig. 1), we implemented a simplified version of Radiance Scaling, by integrating the feedback of cultural heritage professionals during several projects. This implementation is a simple plugin, accessible in the Meshlab menu “Render > Shaders”.

Once selected, our Meshlab Radiance Scaling plugin provides four general options (Fig. 2):

- **“Enable Radiance Scaling”**: activate or deactivate the shape enhancement (Fig. 3).
- **“Invert Effect”**: the rules of concavities and convexities are inverted. This has proven to be useful in certain cases for providing more legible results. In fact, it mimics the work of archaeologists when making squeezes of ancient imprints.
- **“Display mode”**: choose the different shading styles. The different styles are described below.
- **“Enhancement”**: the enhancement factor that can be adjusted between 0 and 1, with 0.5 its default value. A 0-value corresponds to no enhancement, and thus has the similar effect as turning off the shape enhancement.

Among all the shading possibilities with Radiance Scaling, we decided to implement four of them since they are quite intuitive to use:

- **“Lambertian Radiance Scaling”**: the object is rendered with diffuse shading.
- **“Lit Sphere Radiance Scaling”**: a lit sphere (Sloan 2001) encodes the lighting environment and reflective properties into an image of a sphere. For further improving the legibility, a different lit sphere can be used for convex and for concave regions (Fig. 3 left).
- **“Colored Descriptor”**: convexities and concavities are displayed with two distinct colors (blue an red).
- **“Gray Descriptor”**: we simply show the enhancement factor for each position of the object.
4. Case Study: the Epitaph of St. Emilion

During an archaeological study of the medieval Saint Emilion (France), the researchers were facing difficulties to decipher the inscriptions of the Costaulus epitaph, situated in one of the underground monuments in Saint Emilion. For the preparation of the virtual inspection of the engraved stone, a 3D acquisition of the artifact was done with a Faro ScanArm (Fig. 4), by Pascal Mora and Robert Vergnieux (PFT3D, CNRS UPS 3551). The resulting 3D model consists of over 7 million polygons (Fig. 5).

The 14-lines long inscription is situated on a calcareous bloc with a length of 2.1m and a height of 0.7m. The Meshlab Radiance Scaling plugin helped to decipher the most erased parts of the inscription (Figs 6-7), and to reject the hypothesis about the presence of an inscription of the name “Emilion”.

5. Conclusions and Future Work

In this paper, we presented the Radiance Scaling plugin for Meshlab. We have shown that this expressive visualization technique enhances the legibility of acquired 3D models through selective shading, and that the technique is easily accessible thanks to the integration in Meshlab.
In current and future work, we strive to integrate other useful expressive visualization techniques in Meshlab. For instance, we are convinced that the surface relief analysis technique (Ammann 2012) is particularly useful, since details at multiple scales can be analyzed simultaneously.

References


OpenInfRA – Storing and Retrieving Information in a Heterogeneous Documentation System

Alexander Schulze and Frank Henze
Brandenburg University of Technology, Germany

Felix F. Schäfer, Philipp Gerth
German Archaeological Institute, Germany

Frank Schwarzbach
Dresden University of Applied Sciences, Germany

Abstract:
Currently OpenInfRA is being developed as a new system for the documentation and publication of field projects in archaeology. It is based on two legacy solutions and will be realised with OpenSource-technologies as a web application with open interfaces to ease the communication with web services and desktop applications. During the process of developing one research question to be addressed concerns the search across several data types. Due to different methodological approaches the system will have to cope with heterogeneous information in respect to data quality, formats, types, and languages. Thus, search queries and visualisation of the results become a technical and semantical challenge as there is a need for incorporating impreciseness and proximity into a logic-based query language. The approach taken is to integrate special retrieval algorithms that combine the power of Boolean logic with vague and uncertain conditions, implemented with a calculus query language named CQQL.

Keywords:
Database, Data Integration, CQQL, QSQL2, Fieldwork Documentation System

1. Introduction

In recent years more and more research projects (both done by fieldwork and by literature or material studies) of the German Archaeological Institute (DAI) have started to create, process and store their project data and results with digital tools (Schäfer and Thänert 2011). Until recently, there was a strong tendency by new projects to develop their own individual systems regarded as being optimised for unique research questions at specific places. Although some money was spent and a lot of time was invested, these solutions often lacked professional support and a model that would be viable beyond the end of financial funding. These changes and challenges lead to a need for a central, uniform system making use of synergy effects across several research projects and addressing the most common digital tasks in archaeological fieldwork: to document, analyse and archive the data of different researchers and to allow for collaborative work with different disciplines and institutional partners.

As the DAI conducts a high number of diverse research projects, varying in thematic, methodological, spatial and temporal coverage, such a new system – currently named OpenInfRA (= Open Information System for Research in Archaeology) – will have to offer a design and implementation guaranteeing flexibility in order to be customisable for individual needs of specific projects. At the same time, the underlying data model will have to have a common structure and mandatory core elements to insure a minimal standard of data consistency and to enable a preferably high level data-mining across several, independent projects.

An additional requirement of the system concerns its interoperability with other
applications, both with online-resources such as databases, geo-services or library catalogues and local, specialised desktop applications. It should also support long-term preservation, especially by providing extensive and consistent metadata for the raw data stored and a sustainable business model. Many projects of the German Archaeological Institute as well as of departments at different universities are closely linked with researchers and institutions spread around Germany and all over the world. These partners require a web based, multilingual access to the digital information about the project that they are working on.

Section (2) of this paper describes the Legacy Systems on which the concepts and requirements of OpenInfRA are based on. It can be assumed that the overall objectives considered for in OpenInfRA are not only valid in this special project. Therefore, the general project aims will be presented in Section (3). In Section (4) we present a case study explaining the motivation for a complex retrieval system. Section (5) provides an overview of the framework of CQQL and Section (6) introduces the query language QSQL2. The conceptual design of the retrieval system within OpenInfRA is then briefly described in Section (7). Finally, Section 8 contains an outlook on future work.

2. Legacy Systems

OpenInfRA will be based on the technical, conceptual and practical experiences made with two legacy systems which were developed independently from each other at the German Archaeological Institute (DAI) in Berlin “iDAI.field” (http://www.dainst.org/de/project/idaifield?ft=all) and the Brandenburg University of Technology (BTU) in Cottbus “CISAR” (http://www.tu-cottbus.de/cisar/).

The modular field research documentation system iDAI.field (Schäfer 2011) has been in development for five years at the IT-department of the DAI. It is based on the commercial software FileMakerPro. The initial database was designed for the specific needs of the German excavations in Pergamon/Turkey, but quite soon became the documentation system for many other field research projects within the German Archaeological Institute. Apart from the dependency on a proprietary system, the main drawback of this solution is the maintenance of these different project databases, as every project gets its own copy which is based on the latest version of a master version. During the individual setup project-specific changes are made. Rolling-out general changes in the master database for all projects is nearly impossible as it has to be done for each of them individually. Due to the high number of projects using iDAI.field, this kind of maintenance proves insufficient as projects cannot benefit immediately from improvements suggested by new users.

As the complexity of the data model and the functional requirements of the system have grown over the years, the user-friendliness of the graphical user interface has suffered. There is no web interface and sharing the database online is only possible through the commercial Filemaker Server licences hosting of up to 4 projects per server only. Thus, currently 12 projects are sharing their data online through four different servers - an unreasonably ineffective procedure. The use of GIS technology in archaeological projects is increasing and at present iDAI.field does not support storing georeferencing information except for point coordinates.

Despite its drawbacks iDAI.field is in use by over 35 projects from different universities and the German Archaeological Institute, building on the knowledge gathered over several years. A lot of feedback and suggestions from users have helped to develop several optional database modules for documentation of excavations, surveys, building studies, iconographic studies, material studies, restoration work and scientific sampling.
The second system, CISAR (Henze, Lehmann and Langer 2007), has been developed since 2004 as a web based application with OpenSource-tools. In its initial phase, it was primarily designed for the need of the “Domus Severiana”-project on the Palatine in Rome by BTU Cottbus and the Baalbek project in Lebanon by the German Archaeological Institute. Therefore, it had a strong focus on architectural studies and building histories. For the Rome-project the integration of a 3D-viewer was realized so that a user can visualize specific query-results on the database as dynamic 3D-models. Fully integrated WebGIS-functionalities were implemented for the Baalbek-project. Although some more projects use CISAR today, it has two principal disadvantages. One is the data model which is not extensible rendering the adaptation of the system for new projects very difficult and time-consuming. The second one relates to the software implementation needing a major redesign to adopt modern technologies and functional requirements.

The evaluation of the legacy systems revealed that CISAR and iDAI.field are complementary to each other. While CISAR has a modern architecture with a good WebGIS client and 3D visualization, iDAI.field offers a flexible data model and long-term project experience which resulted in a broad diversity of features of the tool. It is therefore intended to merge the advantages of iDAI.field and CISAR into one single state-of-the-art system.

3. Project Aims

The overarching aims of OpenInfRA can be summarized as follows:

- creating a web application to enable an efficient, well-structured and consistent storage and use of primary research data;
- taking into account all methods relevant for archaeological and historical fieldwork, such as excavations, surveys, building documentation, drilling cores, geophysics, surveying, restoration, etc.;
- modelling a homogenous, but flexible data structure with predefined thematic modules which can easily be adapted to the needs of specific research questions;
- supporting multilingualism in order to facilitate international cooperation of scientists;
- allowing differentiated access to the content according to a graded digital-rights-management;
- enabling the early dissemination of primary research data and selected result as online-publications, including the possibility for new users to comment on the content;
- using non-commercial, open and documented technologies for web applications and web services and providing the code of OpenInfRA as Open Source;
- providing both an online-version as well as an offline-version to ensure the usability in places and situation with no internet connectivity, including robust, automated procedures for the synchronisation of data from different versions;
- supporting spatial analysis by WebGIS-tools for the generation, administration and presentation of 2D- and 3D-data;
- integrating a 3D-information system based on web technologies for three-dimensional object geometries to help visualize and query correlations of complex buildings and stratigraphical sequences;
- implementing standardised interfaces both to integrate external resources and web services as well as to disseminate content for online-portals or local applications;
• defining and documenting a XML-scheme which allows for the export and import of data in order to increase its reuseability and its long-term sustainability;

• mapping of the data model to CIDOC-CRM to increase the interoperability of the data;

• realizing different search and retrieval strategies (e.g. for experts or untrained users) and providing filters and facets to narrow down result sets;

• making use of international and national standards and services like SKOSified vocabularies, authority files or gazetteers.

The context diagram of figure 1 gives an overview of the planned interfaces to other existing offline and online systems.

4. Retrieval Motivation

In order to conduct research questions and deliver sensible scientific results, the retrieval of information becomes a major issue in developing OpenInfRA and requires additional technical research. Three characteristics of archaeological information - the impreciseness, the incompleteness and the heterogeneity - are special challenges for relational database systems which have been optimised for defined and complete data. This applies in particular, but not limited to cases where legacy data is mixed with new data in a research project. The documentation of old field projects often lack metadata which is regarded as crucial today, for example excavated objects that have no associated and recorded context. The aspect of imprecise data occurs also in modern, high standard field projects, e.g. when colours of soils are not defined and harmonised and the conditions prevailing when a specific soil is described are not documented (e.g. the differences in colour between dry and wet earth). Furthermore, the structure of the information differs in OpenInfRA as not only tables and attributes contain potentially searchable information, but also attached documents (images, drawings, PDFs, sheets, documents, etc.) as well as 2D and 3D geometries.

Information systems that are based not only on databases, but also on files and documents are extremely useful for saving huge amounts of data. But without corresponding techniques to retrieve this data, the stored information will end up as a final depot. Our intention is to combine the power of a relational database and their query language SQL with the impreciseness and proximity of the information retrieval. To achieve this, we will draw on a framework for uncertain retrieval on certain data that is called CQQL and is implemented as QSQL2. In addition to many possible retrieval methods and filters, we will also try to keep the Graphical User Interface (GUI) as smart as possible.

An archaeological question that exemplifies the complexity of the task would be as follows: “Find all sculptures which have the original height (preserved or reconstructed) of around 1.80 m (e.g. around life-size) and search not only in the current database for stone finds but look also in the old excavations reports which have been attached as PDFs to the excavations datasets.”

2 QSQL2 is an extension of the logic-based query language QSQL (and therefore of SQL). Further information about QSQL can be found in (Lehrack and Schmitt 2010).
5. CQQL

CQQL (Commuting Quantum Query Language) is a calculus query language that was developed by the Chair of Database and Information Systems at the BTU. At this point, we will only give a short introduction of the basic idea behind CQQL and the possibility to use weights. The following paragraphs in this chapter are quotations from longer descriptions in (Lehrack, Schmitt and Saretz 2009, 2-3, 10).

5.1 Basics

In general, CQQL enables the logic-based construction of queries out of traditional Boolean and similarity conditions. The underlying idea is to apply the theory of vector spaces (Salton and Lesk 1986), also known from quantum mechanics (Nielsen and Chuang 2000) and quantum logic, for query processing. Table 1 gives correspondences between query processing concepts and the vector space model of CQQL.

<table>
<thead>
<tr>
<th>query processing</th>
<th>CQQL model</th>
</tr>
</thead>
<tbody>
<tr>
<td>value domain</td>
<td>↔ vector space (H)</td>
</tr>
<tr>
<td>tuple to be queried</td>
<td>↔ vector (v(t))</td>
</tr>
<tr>
<td>condition of query</td>
<td>↔ condition space (vs[c])</td>
</tr>
<tr>
<td>evaluation</td>
<td>↔ squared cosine of the angle between (v(t)) and (vs[c])</td>
</tr>
</tbody>
</table>

Table 1. Correspondences between query processing and the model of CQQL.

5.2 Weighting

Using more than one similarity condition on different attributes often causes a new problem: Which importance on the result should one condition have in comparison to another one? So, the user could decide that the impact (weight) of one condition should be only half of the influence of another condition. This can be realized by assigning a weighting variable out of the interval [0,1] to each condition.

To elucidate the mechanism behind the weighting we will examine two extreme cases in more detail. A weighting variable of 0 leads to a behaviour that the corresponding condition has no longer any effect on the final evaluation result. In contrast to this, if both weight variables are equal to 1, we achieve the same evaluation result generated by applying the unweighted versions of conjunction and disjunction. For further detailed discussion see (Lehrack et al. 2009; Schmitt 2007).

6. The Query Language QSQL2

“The Structured Query Language (SQL) is the de facto standard for accessing relational database systems. Since its introduction in the 70s and its first standardization in 1986 the practical significance of SQL has grown enormously. In this section we want to extend traditional SQL by occurrence and relevance probabilities. For this purpose, we developed a new SQL dialect called QSQL2. For defining syntax and semantics we want to refer the core functionality of SQL-92 (Date 1996) which covers the well-known relational algebra operations: selection, projection, join, union and intersection. Additionally, we present a grouping operation and the application
of nested predicates.” (Lehrack, Saretz and Schmitt 2012).

6.1 Defining Tables

In order to make use of the standard QSQL2 functionalities, no special tables are needed. This means that QSQL2 can be used for any existing database table without any modification. For calculating similarity values, QSQL2 will provide a set of standard scoring functions for each data type and different domains. Any authorized database administrator can add additional user-defined scoring functions. The following definition (Lehrack, Saretz and Schmitt 2012) shows how to create a new table including an optional user-defined scoring function.

```sql
create table table-name (
    column-name data-type [~scoring-function()],
    ... )
```

For example, the fictional table `sculptures` is built by the statement:

S1: `create table sculptures(
    name varchar(50),
    type varchar(50),
    reconstructed_height int ~height_sf(0, 20),
    preserved_height int ~height_sf(0, 20),
    material varchar(50),
    complete boolean )`

The standard syntax for a create table statement (Date 1996) is extended by the following concept:

```sql
~height_sf(0, 20):
```

By this expression we set a specific scoring function for the attribute height. In the statement above we defined a scoring function based on a restricted domain ranging from 0 to 20. Concretely, the score value/similarity between two height values \( h_1 \) and \( h_2 \) is calculated as:

\[
SF(h_1,h_2) = (1 - \frac{1}{20} \cdot |h_1 - h_2|)
\]

For instance, the predicate `preserved_height ~ 1.8` for a height of 3.0 is evaluated to 0.94. QSQL2 supports a set of different scoring functions for each data type and different domains. Additionally, it is possible to integrate user-defined scoring functions by an authorized database administrator (Lehrack et al. 2012).

6.2 Similarity Conditions on Certain Data

“In standard SQL a homogeneous set of tuples forms the outcome of a given query. As already stated, there is a different situation for queries employing similarity conditions. In principle, a query in QSQL2 is formulated in the same way as in standard SQL. Thus, we also use the well-known select-from-where pattern to define (i) an output structure by the select-clause (select-attribute-list), (ii) a data basis by the from-clause (table-subquery-list) and (iii) a logic-based condition by the where-clause (search-condition). The main structure is described by following simplified pattern:” (Lehrack, Saretz and Schmitt 2012)

```sql
select [distinct]select-attribute-list
from table-subquery-list
[where search-condition]
[order by order-attribute-list]
[top top-k-number]
```

As an example, we examine the database part of our fictional archaeological example from chapter 4: “Find all sculptures which have the original height (preserved or reconstructed) of around 1.80 m (e.g. around life-size):”

Q1: `select id, name, preserved_height AS p_height, reconstructed_height AS r_height
    from sculptures
    where preserved_height ~ 1.8`
(i) select-clause:

The column structure of the result table is defined by an attribute list specified in the select-clause. QSQL2 allows duplicated tuples in tables. If you intend to eliminate these duplicates, you have to use the keyword distinct in the select-clause. A list of all attributes can be abbreviated by an asterix (e.g., select*).

The computed score value (attribute escore) is always part of the result table. (Lehrack, Saretz and Schmitt 2012).

(ii) from-clause:

A from-clause contains a list of deterministic tables and/or subqueries. Based on this list a set of tuple combinations is generated by a cartesian product (Elmasri and Navathe 2002). The generated tuple set represents a data basis for an optional selection condition given in the where-clause.” (Lehrack, Saretz and Schmitt 2012).

(iii) where-clause:

To select tuples from the defined data basis (see from-clause) we use a logic-based condition which combines predicates by the logical operators and, or and not. In addition to classical SQL, QSQL2 supports similarity conditions indicated by the new similarity operator (~).”(Lehrack, Saretz and Schmitt 2012).

For our example Q1 we suggest that none of the columns in the deterministic table sculptures has an user-defined scoring function. So a default scoring function for numeric attributes is used to evaluate preserved_height ~ 1.8 and reconstructed_height ~ 1.8. Further, our example table only contains three entries. The result is then given by:

```
Q1: select id, preserved_height AS p_height, reconstructed_height AS r_height from sculptures where preserved_height ~ 1.8 or [0.5, 1] reconstructed_height ~ 1.8
```

With this simple weighting modification we can double the importance of the reconstructed_height in relation to the preserved_height. So our result will change to:

If Q1 was to be executed on a normal database, using SQL, we would get an empty result set.

In chapter 4, another example is mentioned, that could be solved with QSQL2. Describing a specific soil during different conditions, that are not documented, can lead to records of the same soil with different colours (e.g. caused by the colour differences between dry and wet earth). To exacerbate the situation, let’s assume that the colours are

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>p_height</th>
<th>r_height</th>
<th>escore</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Statue of Apollo</td>
<td>1.78</td>
<td>1.9</td>
<td>0.95</td>
</tr>
<tr>
<td>8</td>
<td>Fragment of Zeus</td>
<td>1.6</td>
<td>1.77</td>
<td>0.9</td>
</tr>
<tr>
<td>151</td>
<td>Statue of Augustus</td>
<td>1.65</td>
<td>1.68</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table 2. Result table from Q1.

```
or reconstructed_height ~ 1.8
```

Q2: select id, preserved_height AS p_height, reconstructed_height AS r_height from sculptures where preserved_height ~ 1.8 or [0.5, 1] reconstructed_height ~ 1.8

```
```

---

Table 3. Result table from Q2.

“QSQL2 supports a weighting of subconditions by assigning weighting variable $\omega_1$ and $\omega_2$ to each operand/ subcondition of a conjunction or a disjunction (e.g., syntactically by $\text{and}(\omega_1, \omega_2)$ and $\text{or}(\omega_1, \omega_2)$). A weighting variable $\omega \in [0,1]$ controls the influence of the probability value of the corresponding subcondition on the result (e.g., 0 = no influence, 1 = full influence). The detailed description for evaluation and semantics of weighting subconditions can be found in (Lehrack 2012). A weighted version of query Q1 is formulated by:’” (Lehrack, Saretz and Schmitt 2012)

…”QSQL2 supports a weighting of subconditions by assigning weighting variable $\omega_1$ and $\omega_2$ to each operand/ subcondition of a conjunction or a disjunction (e.g., syntactically by $\text{and}(\omega_1, \omega_2)$ and $\text{or}(\omega_1, \omega_2)$). A weighting variable $\omega \in [0,1]$ controls the influence of the probability value of the corresponding subcondition on the result (e.g., 0 = no influence, 1 = full influence). The detailed description for evaluation and semantics of weighting subconditions can be found in (Lehrack 2012). A weighted version of query Q1 is formulated by:’” (Lehrack, Saretz and Schmitt 2012)

Q2: select id, preserved_height AS p_height, reconstructed_height AS r_height from sculptures where preserved_height ~ 1.8 or [0.5, 1] reconstructed_height ~ 1.8

With this simple weighting modification we can double the importance of the reconstructed_height in relation to the preserved_height. So our result will change to:

If Q1 was to be executed on a normal database, using SQL, we would get an empty result set.

In chapter 4, another example is mentioned, that could be solved with QSQL2. Describing a specific soil during different conditions, that are not documented, can lead to records of the same soil with different colours (e.g. caused by the colour differences between dry and wet earth). To exacerbate the situation, let’s assume that the colours are

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>p_height</th>
<th>r_height</th>
<th>escore</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Fragment of Zeus</td>
<td>1.6</td>
<td>1.77</td>
<td>0.9</td>
</tr>
<tr>
<td>12</td>
<td>Statue of Apollo</td>
<td>1.78</td>
<td>1.9</td>
<td>0.95</td>
</tr>
<tr>
<td>151</td>
<td>Statue of Augustus</td>
<td>1.65</td>
<td>1.68</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table 3. Result table from Q2.
stored in different formats, e.g. one colour is recorded as RGB value, another as HEX code and another one as a colour name. First of all, we need a user-defined scoring function that can transform our three colour descriptions to a consistent format. Afterwards, it will be possible to compare our soil records respective to their colour and return results that contain similar colours.

QSQL2 can combine the strength of relational database systems and their powerful query language with the above mentioned characteristics of archaeological information - the impreciseness, the incompleteness and the heterogeneity. In addition to that fact, QSQL2 also provides a base for the comparability of the results with other retrieval systems that are able to deliver similarity values (e.g. Information Retrieval on different documents or Image Retrieval on image collections). The improved properties of QSQL2 come with a higher need of performance, because some mapping and transformation algorithms need to take place. At this time we can’t give exact information about performance issues.

7. Conceptual Design of the Retrieval System

The planned retrieval system has a long list of requirements that should be fulfilled in the development process. In order to address a wide range of users with it, we plan to integrate different types of retrieval functionalities. Technical these systems will differ by the way they look and how they search on the data. We can divide them into text and SQL based retrieval systems.

7.1 Text Based Retrieval System

The text based retrieval system will hold an inverted index (Heinrich 2008) on the current data that will be used to answer a user request. The database and the documents will be indexed by the system in a defined interval. The system will support different analyzers like stop word filter, automatic language detection, synonyms (based on SKOS⁴), stemmers and the support of different search operators and placeholders. To give the user access to this index we provide two kinds of interfaces that can be used to formulate a request. The first interface is a simple input field called simple search that can be found on every page of OpenInfRA. For researching into more details, the simple search can be extended to an advanced search. In this case, the user will have more visual options for defining a search query or restricting the search domain. Every extension of the request that will be formed in the advanced search can also be formulated in the simple search by using special keywords. This retrieval system will appeal to users that are not familiar with the structure and the content of OpenInfRA.

7.2 SQL Based Retrieval System

The SQL based retrieval system will make use of the introduced QSQL2 language. This retrieval will be addressed to expert users that know the structure and have an idea of how information is stored inside the database. To generate a SQL based request the user needs no knowledge about SQL or QSQL2. At this point we will provide two special interfaces. The first one is called detail search and was already used in the legacy system iDAI.field. This interface uses the same form for query building which will be used for data entry. Furthermore, there will be options such as calling index lists for quick access to possible input or weighting parts of the form. The second one will be a more database-oriented solution and is called expert search. This should work like a construction kit for database queries. The user can take several visualized parts of the database and put them together to build a statement. Furthermore, the generated QSQL2 / SQL statement will be directly modifiable.

For every request in the SQL based system, the user can decide if he wants to make

---

⁴ SKOS (Simple Knowledge Organization System), visit http://www.w3.org/2004/02/skos/ for further information.
use of the uncertain functionality provided by QSQL2. Whether only certain data should be queried, this will be done by standard SQL and thus produce no additional performance overhead.

7.3 Result Views

After processing the retrieval query and gaining the results we will present them in a suitable way. The user will be able to choose between four different views. Each view is first sorted by the relevance of the results that were delivered by the retrieval system and that are displayed as bar charts.

- (i) list-view: The standard view for all results is the list view. This is a simple view that can be compared with the result list delivered by Google. The results cannot be sorted based on another factor (Fig. 2).

- (ii) catalogue-view: This view will show more details of the results but will still be compact. Each project can redefine a specific selection, order and formatting of a fixed number of attributes that will be shown for each dataset. This selection also determines the order of the results. These attributes could be: a picture, an object id or a document name, a title and a long description.

- (iii) table-view: This is a table-like view with attributes in columns and datasets in rows. Users can individually change the range of displayed attributes for each result. It will be possible to add and hide the shown attributes, limited to a minimum of two attributes. It will be possible to order the results list by one of the displayed attributes.

- (iv) thumbnail-view: This view is specialised for results that contain pictures. Results that contain no pictures will be filtered out. Instead of showing text from attributes in a classic list, the results will be shown as tiles. Only one or two textual attributes will be associated with a thumbnail (e.g. short description and photo number). By hovering over or clicking on a picture of a result, the picture will be zoomed and more information will appear.

7.4 Other

For further restriction we will support a filtering mechanism. It will be possible to set several filters while constructing the search request. This will also reduce the time used to retrieve results. It will also be possible to limit the results in the different result views. To make use of the retrieved data from OpenInfRA, the results will be exportable in different formats. The user will have the possibility to choose the data from the result list that will be exported. He will also be able to select the attributes of the chosen data sets will be used.

8. Summary and Outlook

In this paper, we have presented the new documentation system OpenInfRA which is currently under construction and hopefully can be launched as a prototype by end of 2014. We have presented the planned key features in general and its retrieval functionalities in particular. For this purpose, the syntax and semantics of the query language QSQL2 was introduced and the design of our planned
retrieval system discussed. At present, the retrieval part of OpenInfRA will support the language QSQL2 on certain data. In the future, we will have to examine whether the use of probabilistic tables provided by ProQua5 will be useful for improving the retrieval and the user experience. Currently, QSQL2 maps the SQL:92 standard with some restrictions, e.g. the GROUP BY extension is not supported in the context of similarity requests. The development team of QSQL2 focuses on a stable release and will defer the mapping process, because it consumes too much time. Once the first prototype of OpenInfRA will have been implemented and the search and retrieve functionalities have been tested with first users we can report about the advantages and disadvantages of the described approach.

References


5 ProQua stands for „probabilistic and quantum logic-based database system”. Further information about ProQua can be found in (Lehrack, Saretz and Schmitt 2012).
Towards Reverse Engineering Archaeological Artefacts

Vera Moitinho de Almeida and Juan Anton Barceló
Universitat Autònoma de Barcelona, Spain

Abstract:
Our research focuses on the 3D documentation of archaeological materials from the Neolithic lakeside site of La Draga. The main objective is to understand possible relationships between the form and function(s) of archaeological artefacts. We begin by providing an overview of our methodology based on Reverse Engineering procedures. Next, a description of the main guidelines used to 3D scan and data processing of wooden artefacts. Then, we present the use of semi-automated surface feature extraction for use-wear analysis, as well as preliminary computer simulation tests using FEA to study specific physical and mechanical issues, and therefore trying to determine possible functions for these artefacts. In addition, we intend to provide new data, as well as possible explanations of the archaeological record according to what it expects about social dynamics. Furthermore, we aim to reuse and to repurpose these 3D models in conservation monitoring, preservation, digital archives, and other future researches.

Keywords:
3D Scan, Computer Simulation, Finite Element Analysis, Quantitative Data, Reverse Engineering

1. Introduction

The archaeological lakeside site of La Draga is located on the eastern shore of the Banyoles Lake, in Catalonia. It was discovered in 1990 during the construction works of the Olympic channel, and it is the first prehistoric site in a lakeside environment found in the Iberian Peninsula. It is an early Neolithic village (Cardial-ware phase) which dates from the second half of the 6th millennium cal. AD.

Its topographical location – a small peninsula with the appearance of an island – made it easy to defend. The abandonment of this site, after more than a century of settlement, has been associated with a fire that affected part of or the whole village, but other causes are to be considered, such as a decline in resources, social changes, in search for more fertile agricultural land, or a resettlement somewhere on the lakeshore, although none has yet been discovered.

Since 1991, three different areas have been excavated: Sector A, the upper area (water table approximately 70 cm below the archaeological level); Sector B, the lower area, beside the Lake (water table approximately 40 cm above the archaeological level); and Sector C, the Neolithic lakeshore, which is now under water. These variations in the groundwater level have affected the conservation of the Neolithic objects herein found.

One of the other aspects that make this settlement so unique is the vast number and variety of wooden and other vegetable fibres objects found. The overlap between the archaeological level and the water table in Sectors B and C enabled the preservation of the most important collection of organic material finds from this period, such as the remains of large rectangular huts with oak posts (more than one thousand), numerous and various wooden and basketry objects (up to 18 taxa; around 170 items with distinct functional proposals: construction, domestic and personal items, hunting, fishing, defence) and large quantities of cereal grains and animal bones. Hence, this settlement is a very rich source of information and contributes substantially to our knowledge of early Neolithic settlements in the Iberian Peninsula, as well as in the Mediterranean area (Bosch, Chinchilla and Tarrús 2000, 2006, 2011; Tarrús 2008).
2. Archaeological Data

Before proceeding with the technical procedures of data capturing, processing and extraction, it is crucial to define previously what sort of data are archaeologically relevant to solve a specific problem. In other words, beyond documentation, preservation, and visualization issues, to what extent can such data generate useful information and how can we translate it into knowledge? Even when using complex technology as laser scanning and the like, archaeological data remains a passive entity, whose descriptions are so ambiguous that no explanation is possible. In this paper we approach this problem distinguishing data capture from data representation, and introducing the need of archaeological artefacts to act as dynamic entities, whose description should enable researchers and the public to “use” them in the way scientific hypotheses suggest.

It is our view that the real value of archaeological data should come from the ability to be able to extract useful information from them. This is only possible when all relevant information has been captured and coded. However, archaeologists usually tend to only consider very basic physical properties, like size and shape/form. Sometimes, texture, that is, the visual appearance of a surface is also taken into account, or the mineral/chemical composition. The problem is that in most cases, such properties are not rigorously measured and coded. They are applied as subjective adjectives, expressed as verbal descriptions, preventing other people from using the description without having seen the object. As to metadata, there is also a significant lack of references to the object’s structural properties, relevant for technical and functional knowledge on the way the object could have been used in the past. However, there is not yet any formalized proposal for documenting technical and functional properties of archaeological materials.

3. Method

Reverse Engineering (RE) is the process of extracting missing knowledge from anything man-made, by going backwards through its development cycle and analyzing its structure, function and operation (USAITA; Dennet 1991; Eilam 2005; Raja and Fernandes 2008; Wang 2011). It consists of a series of iterative steps, each addressing different questions regarding, in this case, an overall artefact. These steps may be repeated as often as needed until all steps are sufficiently satisfied. In this context, we propose a methodological framework based on RE processes (Fig. 1): from the physical-to-digital stage to the interpretation stage, by simulating the artefacts’ function and inferring possible inherent working processes. Our intention is to be able to understand the functional aspects of archaeological artefacts (Moitinho and Barceló 2011). Archaeological objects must be documented in full using all physical information to determine their possible functions in the past. Therefore, instead of traditional descriptions of archaeological material, we suggest to use geometric models, as generated using 3D scanning and appropriate software, and integrating in them all structural, material and physical properties (solid model).

For the current project, we have selected all the wooden artefacts of the archaeological
lakeside site of La Draga, with hunting, fishing, and defence actions as functional hypotheses. Among them, we should make emphasis on 3 bows (the oldest Neolithic bows in Europe, dated between 5,400 and 5,200 cal. AD), as well as several arrows, darts, a spear, and spearheads (Bosch, Chinchilla and Tarrús 2000, 2006, 2011; Tarrús 2008). Most of the wooden artefacts have been restored - from 1995 to 1997, they were sent to the laboratory of restoration at Laténium, Parc et Musée d’Archéologie de Neuchâtel (Switzerland), to be lyophilized; since 1998, they have been sent to the Centre d’Investigacions Subaquàtiques de Catalunya (CASC) restoration laboratory - and are now deposited at the Museu Arqueològic Comarcal de Banyoles (MACB), Girona.

3.1 Three Dimensional Surface Data Capture

Given the specificities of these artefacts – overall dimensions, type of raw-material, surface/texture (micro-topography generally smooth, but with visible use-wear traces), and as they are very fragile and made of a perishable material – it was important for us to document them with as much detail as possible, to avoid manipulating them further, for cyclic monitoring and preservation, and for future researches.

To first proceed with the capture of the three dimensional geometric digital models and new data concerning to the individual form of each item, we used a non-contact close-range 3D structured light scanner – SmartSCAN3D Duo System, Breukmann –, with the shortest FOV available for this model, the 90 mm set of lenses, which has the highest resolution and gives the maximum level of detail (according to the manufacturer, x, y resolution is 50 µm). It is crucial to have a thorough understanding of the sequential steps, because the final outcome depends intrinsically on all of them. Consequently, each step’s parameters must be specially tailored according to clear objectives previously set.

Nonetheless, there are indeed alterations of the original artefact in size, form, texture, and colour, due to taphonomic and post-excavation factors. To minimize measurement error, a sufficient amount of overlapping area between scans was needed. To overcome most of the self-occlusion problems, we used multiple viewing angles during scans, by changing the scanner’s position in respect to the artefact. However, the entangled geometry of some wood knots invariably led to some small holes in the mesh.

Due to logistic matters and to the short time available, after calibrating the scanner we decided to continue only with the artefacts scanning, more precisely with the point cloud capture – including their pre-alignment and alignment, to ensure that there weren’t any relevant parts of the form missing, as well as the quality of the recorded data – at the MACB, using the scanner’s capturing software Optocat 2009. Inasmuch we are only concerned with the artefact’s geometric data, we didn’t capture any image texture. The scan data cleaning, merging and polygonal mesh generating were done later at IMF-CSIC, Barcelona, using the same software.

3.2 Three Dimensional Surface Data Post-processing

The three dimensional surface data post-processing stage consists in processing the 3D data formerly captured – from scan data...
cleaning, to point clouds final alignment, scans merging, and polygonal mesh generation. At the end of this stage, we aimed to obtain a 3D faceted surface model (Fig. 2), and export it in STL format to carry on with feature extraction.

### 3.3 Three Dimensional Surface Feature Extraction

Hitherto, the insufficiency and lack of a clear consensus on the traditional methods of form description – mostly visual, descriptive, and qualitative – have invariably led to ambiguous and subjective interpretations of its functions. It is thus strongly advisable to systematize, formalize and standardize methods and procedures more objective, precise, mathematical and quantitative, and whenever possible automated (Barceló 2009, Moitinho and Barceló 2011).

This stage took place at the Computer Simulation Lab, UAB, Bellaterra. It consisted of extracting quantitative data from the previous 3D surface models, in a way it could be decoded and understood by the archaeologist. Depending on the type of artefact, different sets of features were extracted. This new information provided meaningful data to distinguish one artefact from another. We used both Rapidform XO Scan 2010 (INUS Technology) and MeshLab V1.3.0 (Visual Computing Lab, ISTI-CNR) softwares – the former software had already been acquired and the latter is an open-source software – to take the opportunity to compare, as end users, some of the tools efficiency and output data. For the most part, MeshLab was able to compute the geometric data and topological measurements necessary for our study. As we weren´t able to open any STL files larger than 400 MB in MeshLab, we overcame this problem by using only Rapidform for the heavier files. Whenever possible, we used the latter software to compute the same features, to crosscheck them, and ensure that the data values were consistent.

Differential geometry is based on the idea of curvatures. Measuring curvature allows analyzing the bending energy of surface irregularities. The value of curvedness is invariant to translation and rotation, and describes how much the surface in the neighbourhood of a considered point deviates from a plane. The plotted colormap is from red to blue, corresponding to high (i.e., convex) to low (i.e., concave) values of the curvedness, whereas grey colour indicates planar patches. Briefly speaking, it allows to detect edges and patterns, and possible use-wear macro traces and working surfaces (Fig. 3) – considering that the manufacturing process and functions affect both certain regions and the overall form of the artefact. Although MeshLab also permits to analyze the amount of curvature at a surface point in a semi-automatic way, we chose to use Rapidform only because its tool seemed more efficient for us.

### 3.4 Computer Simulation

The purpose of documenting archaeological objects is to be able to “use” them in the same way they were used in the
past. Obviously, archaeological objects cannot be used in a real way, because they must be preserved, but we can approach them in a virtual way. In this domain of application, form and size data are indeed necessary, yet clearly insufficient to understand functionality issues. Computer simulation can reveal to be a key aspect of archaeological documentation, because it allows seeing ancient artefacts as dynamic entities and not as passive objects.

Artificial Intelligence techniques, in particular computer simulation, permits us to test different features and replicate distinct behaviours on a specific 3D digital model of an archaeological artefact – here described as a mathematical model that incorporates several variables. That is to say, the use of computer simulation as an experimentation and validation tool towards a better understanding of archaeological artefacts, by endowing 3D digital models with both physical and mechanical properties, and thereafter manipulate virtually these enhanced multidimensional models (Fig. 4) (Reichenbach and Kovačić 2003; Kamat and Martinez 2007; Perros 2009).

Given that we already had the 3D digital faceted surface models, we were then able to convert them into 3D digital solid models, and now proceed to simulate kinematics and to analyze the effects of different forces applied to the object’s model. For this project we have been using Solidworks Simulation Premium 2011 software (Dassault Systèmes), at the Computer Simulation Lab, UAB. This software provides several tools for testing and analyzing the form, material, motion, function, and multi-physics of artefacts, wether they are parts or assemblies, by setting up virtual real-world environments and operating conditions. Before running any type of simulation tests it is necessary to follow a few steps, to ensure best results.

3.4.1 3D Solid Model

Requicha (Requicha 1980) refers that a solid must have an interior, which must be determined unequivocally by the solid’s boundary; the form of any solid model is invariant to location or orientation; applying boolean operations to a solid produces other solids; and that a solid model must have finite aspects, for instance, a finite number of faces.

Whereas a 3D continuous surface model is generated after a 3D faceted surface model, which in turn is based on a polygon mesh, and represents the exterior geometry only; a solid model lies on a volumetric mesh, which describes both the exterior surface and the interior volume of an object. A solid model can be defined in terms of a computer representation, i.e. a digital model, of a physical entity, with computable mathematical properties which allows emulating the physical systems behaviour of the real-world artefacts and processes. This definition is to be true as long as it follows the premises that any constructed representation should (Shapiro 2001):

- Be valid in the sense that it corresponds necessarily to some real physical object;
- Represent unambiguously the corresponding physical object;
- Support, at least in principle, any geometric queries that may be asked of the corresponding physical object.
In other words, solid models emphasize on informational completeness, physical fidelity, and universality of representations.

The objective of this step is, thus, to obtain a 3D digital solid model. It comprises:

- Preparing the surface mesh;
- Creating a 3D continuous surface model;
- Converting this surface model into a solid model (Fig. 5).

In the case of the Neolithic bows found at the La Draga site, this procedure has generated parabolic tetrahedral solid elements, which are second-order or higher-order elements, here defined by four corner nodes, six mid-side nodes, and six curved or straight edges. The generated solid mesh uses the Voronoi-Delaunay meshing technique and depends on three factors: the meshing options of the chosen simulation study, the mesh control specifications (defining the element sizes at different regions in the model), and the contact or connections between models or components.

To continue towards our goal, the solid model has to be subdivided into a finite set of connected elements. The basic concept of Finite Element Method (FEM) lies in that a body or structure may be considered as an assemblage of many smaller cells, typically parabolic tetrahedral solid elements. The original body or structure is then decomposed into finite dimensions, whose elements are connected at a finite number of joints called nodes or nodal points, with determinable degrees of freedom.

Nodes are assigned at a certain density throughout the prehistoric bow depending on the hypothetical stress levels of each particular area. Regions which will receive large amounts of stress are modelled having a higher node density than those which experience little or no stress. Points of interest may consist of fracture point perceived at the archaeological object, corners, complex detail, and high stress areas, among others.

The resulting mesh of finite elements acts like a spider web carrying the material and structural properties for each region in which we have decomposed the object, so that those properties can be formulated and combined to obtain the properties of the entire artefact. Equilibrium equations for the entire artefact are then obtained by combining the equilibrium equation of each element, ensuring the continuity at each node. The necessary boundary conditions are then imposed and the equations of equilibrium are solved to obtain the required values of Stress, Strain, Temperature, Distribution, or Velocity Flow, depending on the functional problem to solve. Additionally, such dynamic decomposition of the prehistoric bow-and-arrow has enabled the analysis of how each node or the whole assembly will react to distinct forces and magnitudes, such as certain stress levels, while indicating the distribution of stress, displacement and

Figure 5. 3D digital solid model of spear (D01-KD89-10). Fixing surface errors (top), and solid model elements (bottom).
potential body deformation. Thus, instead of solving the problem for the entire structure or body in one operation, the attention is mainly devoted to the formulation of properties of the constituent elements. In this way, we increase the prediction accuracy in important or critical areas, by reducing it in others not so relevant functionally speaking. The accuracy of the simulation results is intrinsically linked to the quality of this new finite element model.

Even though the geometry of the model has to be optimized before a simulation can be achieved, the final solid model must carry all the relevant information.

3.4.1 Material Composition

Including mass and assigning the raw-materials’ physical and mechanical properties to each artefact and its components can benefit reasoning about the La Draga’s bows functionality. In fact, these are properties whose values should be included – along with, for example, geometry, texture, colour, weight, or name of the raw-material – whenever describing an artefact. We have created a specific material library for the wooden artefacts of La Draga, where mass density, tensile strength, compressive strength, yield strength, elastic modulus, shear modulus, and Poisson ratio, were specified. Each type of simulation analysis and material model determines which mandatory properties’ values fields must be filled in. For ancient bow-and-arrow using (in hunting and/or battle fighting) these were the mandatory properties carrying out functional information.

Since we weren’t able to find neither existing material libraries with the woods which the artefacts of our study are made of – *Taxus baccata*, *Buxus sempervirens*, *Salix sp*, *Cornus* and *Corylus Avellana* –, nor in the available literature all the required physical and mechanical properties’ quantitative data, we have conducted real-world tests to obtain the values in question. Such tests have been difficult given the nature of *Taxus baccata* as a protected species in the Iberian Peninsula; and the fact that *Buxus sempervirens* appears nowadays as a shrub in open woods, thus making very difficult if not impossible to find samples with the minimum required dimensions for the test. Fortunately, some public institutions managed to arrange us already a few *Salix sp* and *Corylus Avellana* logs, and a couple of logs of *Taxus baccata* are on the way.

Next step consisted in cutting the wood samples according to the ASTM D 5536-94 international standard (ASTM 1994), to thereafter be able to conduct both physical and mechanical tests, at the Department of Mechanical Engineering and Materials, Universitat Politècnica de Valencia. Since the ASTM D143-09 international standard (ASTM 2009) requires bigger samples than the equivalent Spanish standards UNE 56 533-539 79 (AENOR 1979, a, b, c, d, e, f), and some of the wood logs weren’t big enough, we decided for the latter standard.

The fundamental structure of wood, from the molecular to the cellular or anatomical level, determines the properties and behaviour of wood. Because of the fact that this material is heterogeneous and anisotropic, in both its hygroscopic and mechanical behaviours (FPL 1999), it was necessary to perform tests (Fig. 6, Fig. 7) not only parallel but also perpendicular to the wood’s grain.

3.4.2 Tests and Analysis

There are several issues to be considered in order to conduct tests, analyze and predict how the virtual artefact would behave as a physical object in possible scenarios of real world operating conditions, namely: form and dimension of the model, mesh density, contacts and connections between component, material properties, the mechanics of human movement and of an artefact assembly (kinematics), type of medium, and physics.
This step consists of:

a) Selecting the type of simulation (Solidworks 2012), namely:

- Static, which calculates displacements, reaction forces, strains, stresses, and factor of safety distribution;
- Frequency, which calculates stress caused by resonance;
- buckling, which calculates large displacements and failure due to axial loads;
- Fatigue, which calculates the total lifetime, damage, and load factors due to cyclic loading;
- Nonlinear, which calculates displacements, reaction forces, strains, and stresses at incrementally varying levels of loads and restraints;
- Dynamic, which calculates the model’s response due to loads that are applied suddenly or change with time or frequency;
- Motion, which allows defining parameters such as gravity, type of contact and position relationship between components or assemblies.

b) Defining the parameters for the simulation and assigning the parameters’ values and settings.

c) Running the real-time simulation test;

d) Analyzing, comparing and evaluating the output data or checking possible behaviours and functions of the enhanced

*Figure 6. Universal Testing Machine (< 50 kN), UNE 56-537-79 standard – Salix sp yield strength test perpendicular to wood grain (left), Salix sp yield strength graphic (top left), Corylus Avellana yield strength graphic (bottom left).*
multidimensional digital artefact under certain working conditions. If necessary, one can modify the mesh density and other characteristics (FEA), redefining parameters, assigning new values and settings or any other input data, selecting another simulation study or run a new simulation test, to troubleshoot problems or evaluating the validity of the model itself.

Figure 7. Universal Testing Machine (>50 kN), UNE 56-542-88 standard – Salix sp compressive strength test parallel to wood grain (left), Salix sp compressive strength graphic (centre), Corylus Avellana compressive strength graphic (right).

Even though we are still at an early stage of running simulations tests, preliminary results are already providing new insights into the complex dynamics of bow-and-arrow using in Neolithic times. The computer simulates the stress the bow experiments and the moving of the arrow as propelled by the bow, helping us to determine the behaviour of each component by incorporating the effects of force and friction – e.g., ballistic, where the parameters of possible trajectories, elements positions, velocity, acceleration, friction and distance are successively changed and tested. Another simulated action is wood fatigue based on the analysis of the bow behaviour during its estimated life-cycle, as well as the influence of wear in its performance and efficiency, i.e., deterioration by wear, or functional degradation vs. functional robustness.

The simulation of other functional properties and hypothetical behaviours, depending on the archaeological questions and artefacts to be studied, might be more or less suitable, not suitable at all, or evaluated in conjunction with other artefacts and properties.

4. Conclusions

Our approach to document the functional aspects of archaeological objects involves applying RE from the physical-to-digital stage to the interpretation stage, by computer simulating the artefacts’ function and inferring possible inherent working processes. Computer simulation can be understood as an experimentation and validation tool that takes care of many different tasks; as well as a kind of coordinator between the different artefact’s components, properties and behaviours – the archaeological artefact as an enhanced multidimensional model. As a result, we aim to infer the function(s) of artefacts, to reach a better understanding of past social dynamics.

Ideally, the achieved results should be also compared and supported by other types of data, to enable more complete scenarios
Towards Reverse Engineering Archaeological Artefacts
Vera Moitinho de Almeida and Juan Anton Barceló

and therefore an overall understanding of the subject. Moreover, if feasible, one should also conduct real-world testing to completely verify.

Here we present a work in progress, and there is still much work ahead. During this research, it will be also important to analyze and evaluate the potentialities, constraints, quality, robustness and effectiveness of RE processes, by controlling the flow of information and vulnerabilities of the system. At the end, we aim to use these processes in the effort to achieve more efficiently better results, as well as to decrease research time and efforts.

For us, the real value of this research lies on the merge of seemingly unlikely but yet so intertwined scientific domains – an interdisciplinary approach between different branches of archaeology, history of technology, material and mechanical engineering, physics, and software developers, among others.

Aknowledgements

The authors wish to thank: The directors and all members of the archaeological project at La Draga site, and the MACB, for all their support, making possible the wooden artefacts to be scanned. Special thanks to Raquel Piqué, Antoni Palomo, Xavier Terradas, Maria Saña, Josep Tarrús, and Oriol Lòpez. The Institució Milá i Fontanals of the Consejo Superior de Investigacion Científica (IMF-CSIC), for providing the 3D scanner. The Oficina Tècnica de Parcs Naturals de la Diputació de Barcelona; the Parcs i Jardins de Barcelona, from Barcelona City Council; The Natural Park of the Volcanic Area of La Garrotxa, Catalonia; and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), in Montpellier, for managing to arrange us the wood logs. Antonio Nadal Gisbert, from the Departamento de Ingeniería Mecánica y de Materiales del Campus de Alcoy de la Universidad Politécnica de Valencia, for all his work and support on testing the physical and mechanical properties of our wood samples. This research is part of the project PADICAT (Patrimoni Digital Arqueològic de Catalunya), funded by the Obra Social la Caixa and the Asociació d’Universtitats Catalanes (Programa RecerCaixa, RECER2010-05), as well as of the project “Social and environmental transitions: Simulating the Past to understand human behaviour”, funded by the Spanish Ministry for Science and Innovation, under the program CONSOLIDER-INGENIO 2010, CSD2010-00034. This research also benefits from Vera Moitinho’s Ph.D. grant from the Fundação para a Ciência e Tecnologia (FCT), Portugal, co-funded by the European Social Fund.

References


AENOR. 1979d. UNE 56 537 79 – Características físico-mecánicas de la Madera. Determinación de la resistencia a la flexión estática. Madrid: Asociación Española de Normalización y Certificación.

AENOR. 1979e. UNE 56 538 79 – Características físico-mecánicas de la Madera. Determinación de la resistencia a la tracción perpendicular a las fibras. Madrid: Asociación Española de Normalización y Certificación.


Data Analysis, Modelling and Sharing
ARCA: Creating and Integrating Archaeological Databases

Maria del Carmen Moreno Escobar
Pablo de Olavide University, Spain

Abstract:
This article presents the process of design and creation of ARCA, an archaeological database developed within the ATLAS research group. ARCA contains information on over 2200 sites in southern Spain, including their archaeological, chronological and typological characterization, as well as their spatial location, and therefore can be seen as a powerful research tool, accessible via the Internet. Focusing on the different steps followed, and specifically on the decisions leading to the current form and state of ARCA, readers will be provided with an insight into the process of integration and reuse of published and unpublished archaeological data and the different issues faced when integrating different datasets.

Keywords: Database, Archaeological heritage, Internet, Andalusia, Extremadura

1. General Context

Sharing data is becoming increasingly important in a world-wide context. It avoids redundancy when producing information and promotes communication between researchers, although it also requires a certain effort in standardisation and metadata creation. This is true also for Humanities in general and Archaeology in particular, where initiatives such as Open Context (http://opencontext.org/) are trying to develop tools for the publication, dissemination and reuse of data within this discipline. Alongside with these, national and regional inventories of archaeological sites are being created by institutions across Europe, such as the different Sites and Monuments Record (SMR) in England, Wales and Scotland (e.g. West of Scotland Archaeology Service, whose catalogue may be accessed on http://www.wosas.net/search.php), Endovélico in Portugal (e.g. Bugalhão 2002; Vaz et al. 2010) and Patriarche in France (e.g. Cottenceau and Hannois 2002; Chaillou and Thomas 2007), amongst other examples.

Spanish archaeology has not been foreign to these developments, although its scale has been limited to a regional basis due to the political organisation of the Spanish state. Accordingly, a number of databases has been created, featuring Arqueos-SIPHA (developed by Instituto Andaluz de Patrimonio Histórico regarding the Andalusian heritage) (e.g. Fernández Cacho 2002) as an outstanding example which is being lately promoted as Mosaico (Ladrón de Guevara Sánchez 2007; Muñoz Cruz 2007). Despite the potential usefulness of this database, its development has not been continuous through time, experiencing long periods of inactivity, overall in the last decade when the number of archaeological activities carried out in the region increased exponentially due to building and infrastructure developments. Therefore, new information about the Andalusian archaeological heritage is spread over a wide variety of resources, such as publications, private compilations and institutional and public archives. This situation means that every new project needs to spend efforts in review all these data sources. More problematically, it also means that data may be overlooked because of their lack of publication, or limited scope of publication, circumstances which limit their visibility. Wouldn’t it be better if all data were unified and made accessible on the Internet to all researchers? This paper explains how this question has been answered within the Atlas research group of the University of Seville, through the creation of ARCA².

² Acronym of Archivo de Contextos Arqueológicos, in English “Archive of archaeological contexts”: I would like to thank Silvia Fernández...
database integrating a number of datasets concerning the Archaeology of different areas of Southern Spain. At the same time, I emphasise certain issues faced in standardising multiple project databases, and explain how these have been solved.

2. ARCA: Aims and Context of Creation

Having this idea of re-usable and accessible data in mind, the *Atlas* research group (PAIDIS HUM-694) of the University of Seville, in which I participate, decided to integrate all the research group members' data compilations, developed within several research projects, into one unique resource available to all its members. Therefore ARCA was developed with three aims in mind:

- The creation of a sound logical structure able both to store diverse information about archaeological sites across Southern Spain and to admit its expansion with the addition of new datasets,

- To warranty remote access to the database to all registered members, allowing them the edition and query of information contained in ARCA,

- The development of a set of tools designed to query the data and its exportation into formats readable by software like ESRI ArcGIS and Microsoft Office.

Notwithstanding, from the beginning of this project, the design and creation of ARCA and the process of data integration were considered as highly complex given the number of databases within the research group (ca. 10) and the diverse nature of the personal archives and their different organisation and structure. I undertook a pilot study in order to integrate three of these personal archives, bringing together information of more than 2200 archaeological sites. Additionally, the cooperation with a computing technologies company was considered as necessary, due to the requirements of such a database, being SIACROS S.L. ([http://www.siacros.com/](http://www.siacros.com/)) responsible of the implementation and technical development of ARCA as a Web product. Such a step would not have been possible without the funding provided by the Spanish government for the Coordinated Research Project “Comparative analysis of the socio-economic dynamics during the Late Prehistory of Central-Southern Iberian Peninsula (IV to II millennia BC)” (HAR2009-14360-C).

The following sections will describe the procedures and decisions taken during the creation of ARCA.

3. Design and Creation of ARCA

3.1 Conceptual Definition and Other Conditions

One of the first decisions regarding the nature and structure of ARCA was its definition as a database, understood as “any computer-based information system where the data that supports that system may be shared [...] [i.e.] used by a wide variety of applications” (Rolland 1998, 1), in order to allow an efficient organisation and retrieval of the data store within the system. Amongst the different database models, I choose the relational model, which would divide a complex phenomenon into several, simple entities. Each of them are stored in a separated table, while all of these tables are related through a key (Fig. 1). Therefore, relational databases allow (1) to avoid the duplication of data, (2) a higher control over redundancy and inconsistency within the information stored, and (3) to save time and effort when filling and using the database (Rolland 1998, 2).

Regarding the conceptual model of ARCA itself, several examples were considered for a preliminary proposal of its data structure, such

Cacho for her suggestion of this acronym.
as a personal database developed during my Masters dissertation (Moreno Escobar 2009, 42-49), as well as other projects’ databases already created (e.g. Ferrer Albelda 2007) (figures 2 and 3 show the conceptual models of both databases, following the examples developed in Rolland, 1998). From these, a number of entities, attributes and relationships concerning archaeological sites were identified and organised as an entity-relationship model (Fig. 4), representing the foundations of ARCA and which would be further refined during the process of design and creation through the addition of new entities and attributes associated to the storage unit within ARCA: the archaeological site. However, this refinement had to follow various conditions to ensure the database’s efficiency, which are, according to Rolland (1998, 72-74):

- Lack of redundancy, i.e. a piece of data should not be stored in two or more places within the same database, since it produces ambiguity and a certain loss of storage capacity.

- Minimum use of null values, due to the ambiguity ascribed to their meanings, ranging from an absence of associated record to the inaccuracy of the possible values ascribed to the record, amongst other possibilities.

- Prevent data loss, through a careful design of the entities, attributes and relationships within the database.

Along with these technical requirements around the development of the database, a Web application was designed and programmed in PHP language in order to provide with an interface to ARCA, making it accessible to all users through the Internet. However, given the presence of sensitive information within ARCA, in the sense of accurate spatial locations of un-excavated archaeological sites, a limited access to these data was required. This was solved through a security system based on usernames and passwords, which bounded the access to the database to a number of registered members, each one of them having a personal access code.

Another level of security was implemented into ARCA’s design, centring on the privileges of the registered members, through the creation of three different user profiles: Visitor, Researcher, and Administrator. The first one is only allowed to enter the database, view and query the data, whereas the second one also holds rights to add, edit and delete records, in addition to the possibility of exporting query results; lastly, the Administrator apart from
Figure 2. Conceptual model of Antequera Depression MMIX, personal database of my Masters dissertation.

Figure 3. Conceptual model of the sites inventory from the region around the river Corbones (Seville, Spain) (Source: the author, based on Ferrer Albelda, 2007).

Figure 4. ARCA’s preliminary conceptual model. Note the differences in the relationship between chronology and functionality in this model and in figure 2, which intents to straight a semantic mistake in the later.
having the rights ascribed to the other profiles, may also create or delete user accounts.

3.2 Analysis of the Data Sources

Once the first conceptual model of ARCA was defined, the databases whose information was to be integrated into it (from now on, original databases) were analysed for two different reasons: firstly, to identify “new” or unconsidered entities and attributes, in order to avoid data loss, and secondly, to estimate the degree of modification these data would require before their migration into ARCA, given the diversity in nature, classifications and formats employed in all three databases.

The first of the original databases was named *Guadiana Medio DB*, and stored information about archaeological sites located in the river Guadiana valley and occupied mainly during the Copper and Bronze Ages (with a minor emphasis on other Prehistoric and Historic periods), as part of a research project on the Late Prehistory of the province of Badajoz. This database was designed in Microsoft Access 2003 following the relational model, thus consisting in a total of 23 tables, 5 of them defined as look-up tables (i.e. containing the possible values ascribed to a specific field) and the remaining 18 storing information about the sites, ranging from their chronology and typology to the environmental characteristics of their locations. However, metadata was not provided with any of the fields, tables or forms included in the database, while in most cases the records were not standardised in values and formats, and many of the forms and pre-defined queries included in *Guadiana Medio DB* were found to be ill-designed and mal-functioning.

The second of the original databases was *Baeturia DB*, and contained information about the archaeology of *Sierra Morena occidental* (the northern parts of the provinces of Huelva and Seville) during the Late Prehistory (mainly Copper and Bronze Ages, although it also includes some Palaeolithic, Neolithic and Iron Age sites). It follows the relational model of databases, and comprises a set of 17 tables (6 of them being look-up tables), which allows for the storage of information on archaeological sites, bibliography and ceramic items, although only the tables related to sites and bibliography were used. Regarding the documentation about the organisation and structure of the database, there weren’t any reports or metadata ascribed to any of the fields, tables or forms included in the database, although its contents were certainly more standardised than in the *Guadiana Medio DB*.

Lastly, *Antequera Depression MMIX DB* was analysed. This database was designed and developed in Microsoft Access 2003 by the author as a personal relational database of archaeological sites located in the north of the province of Málaga (Spain), taking as data sources two bibliographical compilations carried out by Dr. R. Maura Mijares and Mr. J.R. Menéndez de Luarca and a database about the archaeological heritage of the municipality of Antequera (Malaga) (García Sanjuán et al. 2011). It contained information regarding the identification, spatial location and typological and chronological ascriptions of *ca* 1500 archaeological sites occupied between Palaeolithic and Late Roman times. In contrast with the previous databases, its structure and organisation was documented and metadata was provided for all the database elements (Moreno Escobar 2009, 42-49).

From the previous analyses, a number of action points were derived:

- The addition of new entities and attributes ascribed to archaeological sites, that come from the original databases, in order to provide fuller descriptions and to avoid data loss.
- The need to perform a series of modifications to the data already stored in these databases, regarding issues of formatting, atomicity, and ambiguity of contents.
The establishment of a compromise between standardisation and flexibility in order to promote easier data queries and to be able to describe the sites in detail.

The establishment of correspondences between the data types across all three databases.

The need to keep the traceability of the data included, i.e. the indication of their origin in previous compilations, from field work.

The need of qualifying the information available for each site, regarding its circumstances of discovery, the methodologies employed for its analysis, the amount and quality of the evidence... in order to establish different scales of reliability.

3.3 Extending ARCA’s Preliminary Conceptual Model

From the previous analyses, new entities and attributes were added to the initial conceptual model in order to improve the semantic description of archaeological sites within the database (Fig. 5). These modifications were represented in a new conceptual model (Fig. 5), which presents developments of the initial model (Fig. 4), which are concerned with:

- The sites’ spatial references, including the municipality and province where they are located, a description of how to reach them, the geographical coordinates, and the map from which these were derived.

- The sites’ description, like the presence or absence of fortifications and moats, the level of conservation, the presence or absence, type and number of funerary deposits, the perceived visibility from their locations, and the date of visit. Concerning the site chronology, the existence of radio-carbon dating was also contemplated.

- The interventions carried out in each site, regarding the director or directors, a description of their aims and methodologies and the project’s name, amongst others.
The environmental characterisation of the sites’ location, providing details of the geology, land use, vegetation and hydrology.

The name of the owner of the plot where the site is located.

The bibliography, including attributes such as the specific pages of a publication where a site is located, its physical location (libraries, personal archives, Internet...) and the chronological and geographical scopes of each publication.

Information about the author, scale, title, and creation date of maps photographs.

The traceability of the information, such as the inclusion of the former identification number and the database or compilation at the origin of the data, and descriptive fields for metadata regarding the chronology, typology and the accurate spatial location (through UTM and geographical coordinates), amongst others.

Finally, a set of indexes were included to take account of the differences in quality and accuracy of the data available for each site, inspired on techniques of fuzzy databases (Niccolucci et al. 2001).

<table>
<thead>
<tr>
<th>Section</th>
<th>Sub-section</th>
<th>Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>Identification</td>
<td>Name, Province, Municipality, SIPHA/Arqueos ID, Original database, Original ID, Other names</td>
</tr>
<tr>
<td></td>
<td>Archaeological</td>
<td>Activity, Type of intervention, Year, Director(s), Project’s name, Description</td>
</tr>
<tr>
<td></td>
<td>activities</td>
<td></td>
</tr>
<tr>
<td>Access</td>
<td>-</td>
<td>Access, Owner, Type of ownership</td>
</tr>
<tr>
<td>Description</td>
<td>-</td>
<td>General, Area, Observations, Level of conservation, Comments on conservation, Fortified?, Moat?, Number of burials, Type of burials, Detection method, Chronological index, Methodological index, General assessment index, Comments on general assessment, Visit date, Slope, Visibility to the north, Visibility to the west, Visibility to the south, Visibility to the east, Comments and assessment of visibility</td>
</tr>
<tr>
<td>Archaeological</td>
<td>-</td>
<td>Pottery, Bone items, Lithic items, Metal items, Architectural structures and building material, Other, Distribution</td>
</tr>
<tr>
<td>record</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment</td>
<td>Environment</td>
<td>Geology, Comments on geology, Vegetation, Comments on vegetation, Land uses, Comments on land uses</td>
</tr>
<tr>
<td></td>
<td>Hydrology</td>
<td>Name, Distance, Orientation</td>
</tr>
<tr>
<td>Bibliography</td>
<td>-</td>
<td>Work type (Article, book chapter...), Title, Author(s), Journal/Collective work, Year, City, Editor, Pages where the site is mentioned, Physical location, Abstract, Comments, Topic, General geographic scope, Provincial scope, Chronology</td>
</tr>
<tr>
<td>Spatial localization</td>
<td>UTM coordinates</td>
<td>Point, Description, X, Y, Z, Comments</td>
</tr>
<tr>
<td></td>
<td>Geographic coordinates</td>
<td>Point, Longitude, Latitude, Comments</td>
</tr>
<tr>
<td></td>
<td>Map</td>
<td>Name, Sheet, Scale</td>
</tr>
<tr>
<td>Chronology and</td>
<td>Chronology and</td>
<td>General chronological period, Specific chronological period, General functional typology, Specific functional typology, Doubtful chronological ascription?, Comments</td>
</tr>
<tr>
<td>typology</td>
<td>typology</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Radio-carbon dating</td>
<td>Yes/No?, Comments</td>
</tr>
<tr>
<td></td>
<td>Century</td>
<td>Centuries of occupation, BC/AD, Comments</td>
</tr>
<tr>
<td>Documentation</td>
<td>Graphics</td>
<td>Type, Title, Author, Creation date, Scale, Location, Description, Comments</td>
</tr>
<tr>
<td></td>
<td>Photographs</td>
<td>Title, Location</td>
</tr>
</tbody>
</table>

Table 1. Hierarchic data structure in ARCA.
3.4 Implementation of the Conceptual Model

Once ARCA’s conceptual model was finished, SIACROS S.L. implemented the database into MySQL, an open-source database management system (DBMS) usually employed in the creation of dynamic web sites (Ullman 2012, xiii-xiv). This database was complemented by the development of a Web application programmed in PHP language, in order to provide an interface for registered users (Fig. 6), in order to create, edit, delete, query and export data in ARCA (Fig. 7). In this interface, a proper balance between the number of images and the amount of text was sought, for avoiding effects such as visual fatigue, together with an easy and straightforward organisation of the website and database contents, in order to promote the user-friendliness of the application.

The contents of ARCA were distributed according to a hierarchic structure: a first level constituted of nine sections, which are based on the importance of its contents, is complemented with a secondary division into sub-sections, containing the different fields where the data is actually stored (Table 1, Fig. 7).

Alongside with the technical development of ARCA as a Web-based relational database, a number of modifications were made to the information stored in the original databases, related to issues of (i) data atomicity, (ii) data redundancy, and (iii) standardisation of contents.

Data atomicity: This is one of the requirements of the relational databases, consisting on the capability of a field to store one (and only one) value (Rolland 1998, 41-42), i.e. if a site was occupied during the Copper Age and the Bronze Age, these periods should be stored as two different records associated to the site. This condition was not met in some fields of Guadiana Medio DB y Baeturia DB, which made their individual review and correction necessary.

Data redundancy: This is another of the key aspects of relational databases, stating that a piece of data should not be stored in two or more places within the same database, due to the ambiguity it generates (Rolland 1998, 72). The process of analysis of the original databases indicated the existence of redundancy, firstly amongst attributes present in two or more tables, and secondly, within the content of some fields, where some records stored data that was already made explicit by the definition of the field itself. In both cases, individual revisions and corrections were carried out with additional attention, given the potential risk of data loss.

Standardisation of contents regarding (1) formats, (2) units of measurement and, more importantly, (3) classifications of site
chronology and site typology. The first two sets of modifications were performed (as in the previous cases) through individual revisions and corrections, once standards were set for the fields involved, whereas the third one required a different approach. Regarding the chronological classifications, all variants in the original databases were tabulated and contrasted with the chronological periods listed in the *Tesauro de Patrimonio Histórico* published by *Instituto Andaluz de Patrimonio Histórico*, in order to establish equivalences between them and make possible the standardisation of all records stored in ARCA and related to chronology. Additionally, a new category of generic chronology was established to allow for easier queries of the data (table 2). Concerning typological classifications, all terms employed across the original databases were tabulated and identified with a new set of generic categories in order to establish a double standard: a general classification of functional typologies, complemented with another one, more detailed, which responds to the ascriptions given to each site originally (table 3). This duality allows the classification of a site generally as a “Settlement” and secondly more specifically as a “Villa”, “Oppidum” or “Almunia”, depending on a terminology more suitable to its historic context.

After the completion of these processes of atomization, removal of redundancy and standardisation, the database was given a higher level of stability and consistency, allowing the execution of generic and specific queries about any aspect or quality of the archaeological sites included and recorded within ARCA.

3.5 Migrating Data Into ARCA

Once the database was defined and built, the following step consisted of the data transfer from the original databases into ARCA, process executed by SIACROS S.L. using as guidance a migration document, where all fields from ARCA and the original databases were tabulated as a means of visual representation

<table>
<thead>
<tr>
<th>General chronological period</th>
<th>Specific chronological period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Prehistory</td>
<td>Palaeolithic</td>
</tr>
<tr>
<td>Early Prehistory</td>
<td>Early Palaeolithic</td>
</tr>
<tr>
<td>Early Prehistory</td>
<td>Middle Palaeolithic</td>
</tr>
<tr>
<td>Early Prehistory</td>
<td>Late Palaeolithic</td>
</tr>
<tr>
<td>Early Prehistory</td>
<td>Epipalaeolithic</td>
</tr>
<tr>
<td>Early Prehistory</td>
<td>Mesolithic</td>
</tr>
<tr>
<td>Early Prehistory</td>
<td>Undetermined</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Neolithic</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Early Neolithic</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Middle Neolithic</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Late Neolithic</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Copper Age</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Copper Age (pre-Beaker-Bell)</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Copper Age (Beaker-Bell)</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Early Copper Age</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Medium Copper Age</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Late Copper Age</td>
</tr>
<tr>
<td>Late Prehistory</td>
<td>Transition Copper Age</td>
</tr>
<tr>
<td>Protostory</td>
<td>Bronze Age</td>
</tr>
<tr>
<td>Protostory</td>
<td>Early/Medium Bronze Age</td>
</tr>
<tr>
<td>Protostory</td>
<td>Late/Final Bronze Age</td>
</tr>
<tr>
<td>Protostory</td>
<td>Undetermined</td>
</tr>
<tr>
<td>Roman period</td>
<td>Roman Republic</td>
</tr>
<tr>
<td>Roman period</td>
<td>Early Roman Empire</td>
</tr>
<tr>
<td>Roman period</td>
<td>Late Roman Empire</td>
</tr>
<tr>
<td>Roman period</td>
<td>Late Antiquity / Visigothic period</td>
</tr>
<tr>
<td>Roman period</td>
<td>Undetermined</td>
</tr>
<tr>
<td>Medieval period</td>
<td>Andalusian emirate</td>
</tr>
<tr>
<td>Medieval period</td>
<td>Andalusian caliphate</td>
</tr>
<tr>
<td>Medieval period</td>
<td>Taifas</td>
</tr>
<tr>
<td>Medieval period</td>
<td>Almorávide period</td>
</tr>
<tr>
<td>Medieval period</td>
<td>Almohade period</td>
</tr>
<tr>
<td>Medieval period</td>
<td>Christian medieval period</td>
</tr>
<tr>
<td>Medieval period</td>
<td>Undetermined</td>
</tr>
<tr>
<td>Modern period</td>
<td>Modern period</td>
</tr>
<tr>
<td>Contemporary period</td>
<td>Contemporary period</td>
</tr>
<tr>
<td>Undetermined</td>
<td>Undetermined</td>
</tr>
</tbody>
</table>

*Table 2. General and specific chronological periods employed in ARCA.*
of the correspondences between them (Fig. 8). This document was created alongside the processes of analysis of the original databases, enrichment of the conceptual model and its implementation into MySQL.

3.6 Designing the Query and Export Wizard

After completing the process of migration, the structure and content of ARCA was finished. However, one of the main aims of the database is data retrieval, and therefore, a query and export wizard was defined and built in order to allow searches and transfers of data between the different software employed by the members of the Atlas Research Group (Office, GIS...) (Fig. 9).

The area dedicated to data query and export is divided in two parts: the first one containing the fields that could be searched (depending on the broad and specific chronology, general site typology, keywords and province and/or municipality of location), making possible queries such as “find all sites located in the municipality of Mollina where amphorae type Haltern 7-11 were found”; on the second part it is possible to select which type of data will be extracted from ARCA, following the same distribution of the information in sections intimated in table 1. Then, the export would take place into Microsoft Excel 2003 format.

Complementarily, another option made available in ARCA is the export of the whole database into MS Excel, in order to allow the

<table>
<thead>
<tr>
<th>General functional typology</th>
<th>Specific functional typology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guadiana Medio DB</td>
<td>Baeturia DB</td>
</tr>
<tr>
<td>Rock art</td>
<td>Ritual (rock art)</td>
</tr>
<tr>
<td>Settlement</td>
<td>Settlement, Settlement (?)</td>
</tr>
<tr>
<td>Settlement + Burials</td>
<td>-</td>
</tr>
<tr>
<td>Funerary structures</td>
<td>Funerary structures</td>
</tr>
<tr>
<td>Religious structures and buildings</td>
<td>Religious buildings</td>
</tr>
<tr>
<td>Fortifications</td>
<td>Fortifications</td>
</tr>
<tr>
<td>Undetermined</td>
<td>Undetermined</td>
</tr>
<tr>
<td>Infrastructures</td>
<td>-</td>
</tr>
<tr>
<td>Megalith</td>
<td>Megalith</td>
</tr>
<tr>
<td>Mines and quarries</td>
<td>-</td>
</tr>
<tr>
<td>Lithic workshop and tools</td>
<td>Lithic workshop and tools</td>
</tr>
<tr>
<td>Areas of activity</td>
<td>-</td>
</tr>
<tr>
<td>Catchment area</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3. General and specific functional typologies applied to sites in ARCA.
researchers the maximum flexibility when performing queries in any of the database management systems available.

4. Conclusions: Advantages, Shortcomings and Further Work

The previous sections have reviewed the process of conceptual design and technical development of ARCA as a database integrating information about archaeological sites from three specific areas of Southern Spain. However, its design and structure provide the means to enlarge its basis and records, including new sites and new geographical areas. Both circumstances make ARCA an outstanding tool for researchers interested in the Andalusian archaeological heritage, which is currently employed in my doctoral research as a fundamental resource and in the project “Comparative analysis of the socio-economic dynamics during the Late Prehistory of Central-Southern Iberian Peninsula (IV to II millennia BC)”, amongst others. Some of its strong points are:

- It compiles data characterised by their usefulness, accuracy and diversity, derived both from bibliographic reviews and field survey, stored in a relational database following standardised categories and thesauri with a long tradition,
- It provides an easier access and export of this information, enhancing the possibility of data re-use within new research projects about the archaeological heritage in the South of Spain,
- It allows the continuous updating of the data, establishing links to the resources employed regardless of their nature (field work, research projects, bibliographic references...).
Despite its advantages and potential use, some weaknesses and opportunities for further development have been detected:

- It is necessary to optimise the structure and storage of the entity *Bibliography*, due to a high level of redundancy in the table dedicated to the bibliographical references,

- Regarding the query and export wizard, it would be advisable to provide a preview of the selected data, as well as other exporting format directly compatible with Microsoft Access, in order to avoid performing the required pre-processing to the exported tables,

- The absence of spatial modules allowing the preview of the sites spatial location,

- The restriction on its use to researchers within the Atlas group.

It would be advisable to implement accepted standards for a better description of ARCA’s organisation and content, such as the CIDOC CRM.

Despite being currently employed as data source, the awareness of the strengths and weaknesses of ARCA leads me to consider this database as a tool in continuous development. Notwithstanding the previous, it is fundamental to note that, in its limited scope, ARCA has achieved one of the main goals stated at the beginning of this project, *i.e.* enhancing and promoting the reuse of archaeological data. My hope is that this initiative will not stand alone for a long time in the Andalusian archaeology, and that ARCA will serve as an incentive for increasing the re-use of the information already produced during the last decades of archaeological activity carried out in the South of Spain.

### References


A Database for Radiocarbon Dates. Some Methodological and Theoretical Issues about its Implementation

Igor Bogdanović, Juan Antonio Barceló and Giacomo Capuzzo
Universitat Autònoma de Barcelona, Spain

Abstract:
Databases constitute a fundamental element for archaeological heritage management. Our project pretends to examine analytical techniques based on technologies of artificial intelligence, Bayesian statistics to link different archaeological databases and integrate in one operative system data disseminated in hundreds of different files and servers. As a case study we use the recent referential data base of radiometric dates for prehistory of north-east of Iberian Peninsula. In the case of dating our most remote past archaeologists should associate analytical samples dated by C14, with archaeological materials, some of them dated on stylistic grounds, and with contexts archaeologically excavated and ordered stratigraphically. Our project addresses the concept of “telearchaeology” as the essential condition for conversion of archaeological heritage from mute “stones” into what it should be: a basis for knowledge.

KeyWords:
Radiocarbon, Database, Bayesian Statistics, Knowledgebase

1. Introduction

Archaeological objects by themselves do not make the past and do not transmit anything. In order to “make them speak about the past”, knowledge about the proper history of such an item (Appadurai 1986), i.e. knowledge about all its conditions and aspects related to it, is indispensable. Knowledge about heritage stands even when the object itself doesn’t exist (as in intangible heritage), or doesn’t exist anymore.

The location of artefacts in space and time is a starting point for building the network of relationships between observable remains of human activity in the past. The main pillar for interpretation of past social actions is a definition of chronological framework, since it is impossible to understand the history without knowing the temporal and causal flux of different events. To relate bits of archaeological record and make a congruent history out of it, we need one complex and multidisciplinary architecture of analytical procedures, based on computer technology and statistical analysis tools and methods. Relating information on a right way we can generate knowledge about questions of our interest.

Information is not knowledge but knowledge is impossible without information. Building of archaeological knowledge requires ordered and structured information from many sources. We approach this task starting from the process of data construction and preservation according to the concept of Telearchaeology (Bogdanović 2002; Barceló, Bogdanović, Vicente 2004; Barceló, Bogdanović, Piqué 2004), which is conceived as a set of computational concepts, solutions and tools for connection, integration and manipulation of heterogeneous data distributed on the World Wide Web, and which is aimed to generate archaeological knowledge in a self-sufficient way. In this widely open cloud of archaeological knowledge, the idea of “publishing” and knowledge construction gets a new sense. Telearchaeology is an open framework to discuss the publication of data, theories and interpretations in the most efficient way through the formalization and systematization of datasets, and by using interoperability drivers and connections to different databases or other contents around
the WEB. Asking questions to Telearchaeology generates new knowledge which is stored back to the system and makes it growing and learning permanently.

2. How to Date History: from Archaeological Information to Historical Knowledge

A historical period is an interval of time within which an undetermined number of single events happened. Such particular “historical” events should be understood in terms of the occurrence of social actions that were performed by someone who produced something somewhere and some-when. In general, the duration of a single historical period can be estimated in terms of the temporal extent of performed social actions (historical events). But a mere aggregation of particular and partial events would be misleading if the particular relationship between actions, agents and products is not taken into account. Each partial event was performed in a particular sequence in relation to other partial events, and each one had a particular duration. Duration (“running time”, “lifespan”) can be defined in terms of the difference between two consecutive points within the same trajectory (Fig. 1). Such a trajectory is configured by the particular sequence relating the particular events. To know the particular order of an occurring social action within the sequence we should measure the temporal distance (according to a referential duration, for instance the time needed by the Earth to complete an orbit around the Sun, i.e., one year) between such particular event and a referential event (for instance, today). We call calendar date this measure. Consequently, a simple addition of calendar dates of particular events within a single period do not produce a consequent image of the time interval because of the influence of the possible overlapping of different particular events, and overlapping of different, although related trajectories. Of particular importance is the determination of the starting and final point of the historical period. We need to distinguish a particular discontinuity in the social actions that took place before and after those actions within the period.

2.1 Isotopic Event as a Starting Point of Chronological Inference Chain

Historical events are not observable entities but the sum of actions produced in the past. What we can discover and observe are some material consequences of social actions that happened in the past. We call archaeological events to such phenomena, and such archaeological events are a palimpsest
of lower-level events: the particular action that generated the location of such item at this particular place and moment. We call depositional event to this individual facts. One archaeological event can be configured by many different individual depositional events, with different calendar dates and different durations. Nor the calendar date nor the duration of a depositional event can be physically measured. Instead, we should distinguish the possible occurrence of an isotopically determinable death event. It is the particular moment in which a living being -animal or plant- ceased to interact with the atmosphere and biosphere. We assume that the most probable calendar date of a depositional event will be the nearest possible to the isotopically measured calendar date of the isotopic event, with a standard error determined by the duration of the depositional event (Fig. 2).

Therefore, we should relate each isotopic event with corresponding depositional events, i.e. stratigraphic and taphonomic information of each dated sample. Defining context reliability is a fundamental step for obtaining a true relation between the radiocarbon probability intervals and the depositional event we are referring to. Nevertheless, calendar dates of isotopic events are not enough for building historical chronologies. A particular logical connection should be found within the isotopically determined calendar dates of all determinable death events within the same depositional event. The estimated calendar date and duration of all synchronous depositional events within the same archaeological event will be used to measure the date and duration of events higher in the hierarchy. The calculated calendar date and duration of all archaeological events within a single historical event should be used to compute an estimation of the initial and final position of events within the historical period.

To date history we propose following inference chain:

Isotopic event -> Depositional event -> Archaeological event -> Social event.

2.2 Historical Periods as a Series of Causal Events

Many times there is not enough information to define archaeological events from the description of temporally asynchronous depositional events. Although the relevance of spatial information has been argued early in the history of archaeology, main efforts were faithful to systematization of vertical disposition of layers and objects in order to establish relative chronologies. The key developments for spatial and temporal analysis in the method of archaeological excavation and recording are certainly the Kenyon method of phasing and the Harris principles of stratigraphy. The phasing method, as well as the stratigraphy method concern strata as packages distinguished by sediment homogeneity and specific content. Each individual deposition episode is represented as a node in the graph, and relative chronological relations are shown as lines between the nodes.

Although events which have produced formation of layers have certain duration, the nodes in a Harris Matrix are points in a one-dimensional partial ordering, rather than time spans. Nodes organized in graph by the law of superposition may only describe three situations:

- something is later than, or earlier than something else,
- there are no relationships between two,
- the two are contemporaneous.

On these assumptions Holst (2004) suggest that an accurate structural analysis of chronological consequences of different depositional events, can give us starting and end points of one event. Relating stratigraphic units in that way may not respond only to the
law of superposition, but it may represent their causal relation. The possibility of representation of durations has to be represented by a new concept: “broadly contemporary”, which expand classification of chronological consequences.

Further developments of temporal reasoning have opened the door to new background knowledge for building chronologies. It has been suggested that events happen within “coherence volumes”, where all living and dead participants “meet” (Doerr et al. 2004). Causal relationships and event order information produce a temporal network, which in combination with absolute dates can fix “floating” events in one relative chronology.

History is not a simple succession of episodes on a timeline. It is a flow of events that origin consequences in other, posterior events. Spatial information cannot be transformed into temporal information in any simple nor formal way. We need additional information. This statement has been asserted in modern causal analysis (Shafer 1996, Pearl 2000, Sloman 2005), when it has been formally proved that to connect causally an event with another in the same historical trajectory, four conditions are necessary:

- one event should precede the next one in the trajectory, or be contemporary,
- when two events are independent, there should be a location in the graph representing the historical trajectory where both have their probabilities altered,
- one event tracks a second when the probability of the second is the same in any two graph locations where the first happens and the same in any two locations in the graph where the first fails,
- one event is a positive sign of a second if the probability of the second goes up whenever the probability of the first goes up, and goes down whenever the probability of the first goes down. The probability of the second is allowed to change arbitrarily when the probability of the first does not change at all.

The first condition can be archaeologically tested in a Harris/Holst diagram showing seriated temporal units, although without causal relation between them. The other conditions are much more difficult to analyze in the archaeological record. They are, however, necessary to find an estimation of calendar
dates and duration of historical periods on the basis of their constituting events.

As an ordered sequence of related events, the specific chronology of the historical period should be calculated not only in terms of the chronology of the constituting historical events, but also in terms of the specific order or relationship between them. The chronological order depends on the specific ordering based on causal relations between related events within the same period.

3. How to Formalize History?

We can represent an historical timeline by a formal model of historical events and conjunctures using probabilistic graphs with arrows joining linked nodes (Fig. 3). Nodes are connected up with a directed link, representing the relationship between an event and those causally related to it. Links contain information about the dependencies between the historical events in the form of conditional probabilities. The direction of the link arrows suggests that the nodes higher up in the diagram tend to influence those below rather than, or, at least, more so than the other way around. The basis for this assimilation of “causality” and “conditional probability” lies in the so called “Markov condition” (Hausman and Woodward 1999), which states that any node should be conditionally independent of its non-descendents, given its parents, because a phenomenon must be independent of its non-effects, given its direct causes.

To represent the historical phenomenon of Bronze/Iron Age transition in Western Europe (1000-750 BC), we can suggest historical factors like “Migration continues”, “Relationship with Locals”, “Resource Availability”, (as in example shown in Fig. 3) etc. Such factors can also be seen as decisions to be taken by social agents that lived at those times. Entering evidence into a decision node does not really correspond to discovering information or making an observation, but rather it implies a decision choice. The decision network as a whole represents a decision that historical social agents faced. For example, late-bronze age migrants should take a lot of social and economic decisions if they wanted to move to another place. In this simplified example, the decision may be to consider locals as “enemies”, “neutrals” or “cooperants”. In some sense, their decision will take some direction depending on the amount of weapons and defensive works they see among their local neighbours. Yet, this decision is to be introduced by archaeologist, and to do that, archaeologist needs knowledge. The problem is in the recognized limitation on knowledge processing which can be done in mind of one scientist and that’s why decision making would be better performed as an automated operation.

Links that go into a decision node have a special meaning and are called informational links. They indicate what will be known at the time the decision is to be made. That is, the historian will know the values of all the nodes which have links into that decision node, and she will not know the values of any other nodes. When there are some links entering a decision node, the model should find a decision value for each possible configuration of values of the parent nodes. When constructing a decision net, we do not need to concern with all possible links, as long as we put enough to create a directed path running through all the decision nodes (to indicate their time sequence). Having more relevant information for decision making, always leads to a policy with greater expected utility. Many of the decision links may not be
relevant; they provide information about the past that does not help in the new decision. If these links are included, the optimal decision functions will simply ignore them, and the expected value of the policy will be the same, but they may result in very large relation tables for the later decisions. The fact that conditional probabilities relate factors and responses suggests that some responses will tend to occur more frequently when some factors are present. Therefore, to define the causal relations that take part in the development of a historical event we need to introduce a set of probabilities, one for each of the possible states of each node.

3.1 Defining Conditional Probability

The most basic and straightforward way to define a conditional probability between a node and its parents is to explicitly define what is termed the Conditional Probability Table (CPT) (Fig. 4). The CPT is simply a table that has one probability for every possible combination of causally connected historical, archaeological and/or depositional events. This is an \( N+1 \) dimensional table, where \( N \) is the number of parents.

The set of all evidences entered into the nodes of the model is referred to as a case. A case provides some information about a particular historical, archaeological or depositional event. Case files may consist of many cases (acting as a database, in which each case is a database record). For each case in the file, a computer program implementing our model reads the case and enters it as findings into a causal model. Then probability updating takes place to find probabilities for all the nodes that didn’t have findings. Finally the results are written to an output file.

Conditional Probability Tables can be entered by hand, based on some theoretical assumptions, or they can be learnt automatically using the appropriate algorithm based on Artificial Intelligence techniques. In the same way, we can also use a case file to test the performance of the model, finding error rates, log loss, etc., which is equivalent as testing an historical model with archaeological data. The idea is to pass through a file of known cases applying probabilistic inference to each case to generate new information about it, and then saving the case with the additional information to a new case file. If every ethnographical situation supplies a value with certainty for each of the variables, then the learning process is greatly simplified. If not, there are varying degrees of partial information.

Within the model, single evidence is the value for one of the nodes when it is applied to a particular situation. Positive evidence is knowledge that some social factor definitely appears in a definite state. However, we may know that the value of a node is not at some definite state without knowing what its value is. This is called a negative finding. A third type of evidence is likelihood evidence. In this case we receive uncertain information about the value of some discrete node.

3.2 Bayesian Backbone

In order to discover specific events in archaeological record, and to establish their historical ordering and the relations of causality among them, we propose a specific computational architecture made of databases, ontologies and operators. The backbone of this architecture is the Bayesian approach.

Bayesian learning is the process of automatically determining a representative acyclic graph given data in the form of evidence. Each case represents an example, event or situation that exists or has occurred presumably in the world, and the case supplies values for a set of variables which describes the event. Each variable will become a node in the learned net, and the possible values of that variable will become the node’s states. Structure learning determines the dependence and independence of historical, archaeological, depositional and/or isotopic events and suggests a direction
of causation, in other words, the placement of the links in the acyclic graph expressing the historical period. **Parameter learning** determines the conditional probability table (CPT) at each node, given the link structures and the data. Structure learning can be implemented by incorporating a function which searches through some candidate link structures that are plausible and practical in the study domain, perhaps also adding trial latent variables. For each structure we may use parameter learning functions, and then test the resulting model using additional evidence data. The model that scores the highest (perhaps penalized for complexity) is the best structure.

This is not the proper place for discussing the algorithmic nature of structural and parameter learning (see Neapolitan 2003, Darwiche 2009, Koller and Friedman 2009). Finding a maximum likelihood Bayesian network, which is the net that is the most likely given the data is not always an easy task, and many times an interactive approach in which the historian suggest possible links between events, and then tests statistically the direction of causality, provides better results: before we have seen any archaeological data, our prior opinions about what the true historical ordering of archaeological, depositional and isotopic events might be expressed in a probability distribution over the graph of hypothetical causal connections. After we look at the data (and radiocarbon estimates), our revised opinions are captured by a posterior distribution over event relationships.

4. **Where does Information Come from?**

As we pointed out before, historical analysis is based on information and previous knowledge. This information comes from different sources and, even when highest standards in data recording and management are applied, has heterogeneous structure and form. Heterogeneity of data by itself is not intrinsically problematic, but it is challenging for any analysis which needs integrated datasets. Although applied computer techniques solve one part of data recording and management problems, on the other side often open plenty of new ones. Introducing the variety of data without considering data and knowledge generated out of the archaeological primary intervention is one example. Physical, chemical, mineralogical analyses and historic interpretations of individualized archaeological items are usually dispersed in separated and incoherent files and servers. Field documentation databases use to be split in different excavation records and do not integrate data obtained in the laboratory, whose results are registered in detached databases. It is not atypical that data from the archaeological record are badly structured, inadequately formalized and stored in completely isolated datasets, which serve only to reply digitally the old drawers filled with hard-copy archaeological documentation files.

Three main kinds of information – isotopic measures, spatial coordinates, and data about material reality – are necessary for estimating the causal flow of historical events. As a product of different phases and spheres of archaeological research, this heterogeneous information is almost never integrated into a single knowledge base. When we have isotopic estimations, we do not have access to spatial coordinates; when archaeological objects are properly described, we do not have information of the spatial relationships they may have among them. The usually considered small subset of data mostly has a form of passive, qualitative and subjective description, without structural properties or factors that explain formal characteristic of datasets. Information about temporal location of objects is usually given in terms of archaeological phases or “cultural periods”, i.e. in terms of larger or shorter time packages, loosely defined by typology of ceramic or other archaeological material.

The problems raised by bad quality of data, mainly produced during the earlier history of archaeological research on one side,
and heterogeneity of hi standard datasets on another, challenge the telearchaeology approach, and are addressed by our case studies.

4.1 Case Studies

We have built an integrated database for the prehistory of Catalonia containing isotopic events (radiocarbon measures on charcoal, bone and/or seeds) linked to their corresponding depositional events (spatial coordinates of radiocarbon samples, spatial relationships with archaeological artefacts) and in many cases with the corresponding archaeological events (stratigraphical relationships). A preliminary hypothetical chrono-cultural sequence (historical periods) for the prehistory of Catalonia has been created by statistical analysis of actual dataset. The data proceed from some 157 sites, which represent one dated site per 200 km². For the moment, regional availability of radiometric dates is irregular. Dates obtained by standard method on long-lived samples are dominant, while AMS short-lived samples represent barely one third of the dataset. These, and more defects of available dataset are recognized, listed and published (Barceló 2008).

The schema of our initial database was imposed by the structure of available data. Each isotopic event is followed by spatial information and some archaeological observations. Spatial information is given as a position of the site, using actual nomenclature for administrative units (municipality, county, region, province, geographic region), and geographic coordinates (UTM). As the database of isotopic events should serve equally for regional or in-site analysis, here is almost regularly missing important information about coordinates or position of the sample within the site. In fact, the information about depositional event is poor. It is limited to recall material and precedence of the sample, its functional context, type object, and general chronology of correspondent strata. With those data alone we cannot date “history”.

In order to create a causal graph of the historical trajectories of archaeological events, we have begun with a single case, the first agriculture in Catalonia at the beginning of 6th Millenium BC. Available data, in general, do not allow the kind of analysis we have here suggested. Therefore we have limited our investigation to the modern excavation of the early Neolithic pile dwelling site of La Draga on the Banyoles lake. The same archaeological site was excavated in several campaigns from 1990 onwards (Bosch, Chinchila i Tarrus 2000, 2011]. Twenty-four isotopic events have been distinguished, and radiocarbon measured, whose corresponding depositional events have been well documented (Bodganović, Piqué 2012). La Draga site is extraordinary well preserved, and depositional and archaeological events can be defined rigorously, even if the relative ordering of depositional units is stratigraphically complicated given the active post deposition processes.

Additionally, we have investigated the case of the transition between Bronze Age and Iron Age in Western Europe through the examination of more than 1500 isotopic events and their corresponding depositional and archaeological events from the Danube to the Ebro and Arno river. The time-span includes all isotopic events between 1800 and 750 BC., focusing on the last part of the range. The last term is due to the “Hallstatt disaster”. This case study, adopting a macro scale point of view, manages to analyze different phenomena of adoption of technical and cultural innovation which took place in Protohistoric Europe. In particular we aim to describe the diffusion in space and time of different social events cremation burials, new typologies of pottery and metallurgy, the beginnings of complex settlement and fortification, the emergence of “aristocracies”, etc. Our challenge has been to collect lots of information about isotopic events and related depositional events, which were dispersed in different journals and monographs. The limitation of this kind of approaches lies on the critical analysis of available archaeological
information, specially the connection between isotopic measures, isotopic events and their corresponding social facts in the past and resulting depositional events. To tackle this issue we are doing a pre-screening of the samples, giving priority to short-lived samples from closed archaeological contexts or secure stratigraphic layer (Boaretto 2009; van der Plicht, Bruins, Nijboer 2009).

To describe the context we used functional and economic variables in order to provide more information about the society that created the analyzed archaeological record. Therefore, we marked the presence/absence of pottery typologies with a macroscale distribution. We also took into account the emergency of social and economic factors which give us information about the subsistence base through the predominant domestic animal, presence of agriculture, the settlement structure, the construction of some kind of fortification and the exchange networks.

Development of a database structure, building of a model for integration of heterogeneous data stored in different formats and structuring of a system for information discovery, retrieval and representation are principal lines of our case studies. In this research we underline importance of data formalisation and structuring, but with the same attention we consider problems of information accessibility.

5. Conclusions

Dating history is not a simple process of accumulating contemporaneous events. In fact, as we stated before, historical events are not observable entities, but the observable consequences of past social actions which can be archaeologically tested. We understand that historical events are part of a temporal network, and that they always occur as a consequence of past events. Therefore, we must be aware of their causal relations, which are not obvious and implicit in archaeological record. Those kinds of relations can be found in data patterns using computational techniques, and expressed trough probability graph, or other forms of semantic presentation of knowledge. In order to enable intelligent processing, we need complex data structures coming from all possible sources, as well as computer architecture developed on techniques of Artificial Intelligence. The computational simulation of chronological thinking, based on Bayesian statistics, enable an insight in significant causal relations in our datasets.

But archaeological information needed to build historical periods has heterogeneous structure and nature. Up to now our research is limited to two cases of study – a microscale analysis of events within one prehistoric site, and a macro-scale analysis of expansion of innovation phenomena in one part of prehistoric Europe. In both cases we define the historical period as a Bayesian association of isotopic, depositional and archaeological events. We assume prehistory is composed of different layers of empirical evidence: isotopic, depositional, archaeological, social. An archaeological database should be organized in terms of radiocarbon samples (isotopic events) connected to depositional events (spatial coordinates), archaeological events (artefacts) and social facts (hypotheses of action in the past). Therefore, our database has a particular structure in which radiocarbon estimates in one
file are linked to other kind of archaeological data which may be stored at different servers.

Now more than ever we have computational and communicational tools to build an intelligent system which will be able to operate with huge datasets. In the introduction to this paper, we have referred to Telearchaeology as a computer set of concepts, solutions and tools for integration and manipulation of heterogeneous data distributed on the World Wide Web, aimed to generate archaeological knowledge. Based on this concept and following the problem-solving theory, the “DIA-spora” (Distributive Interactive Archaeology – Synchronised platforms of research activities) is a computer architecture that reproduces general framework for archaeological reasoning and data processing and integrate both into a single system (Fig. 5).

Archaeological knowledge has to be distributive, which means, it should come from many different sources, and it should be also interactive, because all scientists should be able to transform collective knowledge. Knowledge and data produced and transformed by individual scientists should be accessible to the rest of the research community in the real time. The result is opposite of traditional fragmentation and built-in obsolesce of knowledge: knowledge use is synchronised in such a way that it is always up to date. Research activities are classified on platforms which hold assumptions and formal requirements for exchange data, hypothesis and information, as well as set of linked databases, and explanatory rules.

If we define Knowledge as Data in a Problem-Solving Framework, the first of two main components of DIA-spora should be a body of explicit knowledge and the second one is a set of actions to manage that body of knowledge. Both components have to be defined in computational terms, that is to say: data structures and procedural instructions.

This theoretical and practical framework offers possibility for collective construction of knowledge which has a potential to accelerate all research activities, but requires certain change in approach to the research and, not less important, in its normative background.

Acknowledgements

This research receives funding from the Obra Social la Caixa and Asociació d’Universitats Catalanes (Program RecerCaixa, RECER2010-05), and partially also by the project HAR2009-12258 by Ministerio de Ciencia e innovación. We acknowledge the collaboration from the research team working at La Draga site, whose work, and partially also ours, is funded by the Ministerio de Ciencia e investigación through projects HAR2009-13494-C02-01 and HAR2009-13494-C02-02. Excavations at La Draga counted with support of Departament de Cultura de la Generalitat de Catalunya, Museu Arqueològic Comarcal de Banyoles and Museu d’Arqueologia de Catalunya. Giacomo Capuzzo acknowledges his research grant from Generalitat de Catalunya.

References


Bogdanović, I. 2002. Internet, el conocimiento distribuido y la resolución del problema arqueológico. MA Diss. Universitat Autònoma de Barcelona.


Standardised Vocabulary in Archaeological Databases

Matthias Lang, Geoff Carver
Georg-August-Universität Göttingen, Germany

Stefan Printz
Hochschule Bochum, Germany

Abstract:
The following paper discusses the use of SKOS-XML vocabularies in database structures intended for documenting archaeological field projects. Following a short introduction on the basics of SKOS, we will outline a tool developed to gather and administrate SKOS-Thesauri from scratch. Then will be shown how we implemented a SKOS-thesaurus in an archaeological database and how we make the vocabularies usable by archaeologists who are not IT-specialists. We will also demonstrate how we use a SKOS-thesaurus to control the language of our database interface in order to offer the participating projects an easy-to-use multilingual interface. We conclude with a discussion of our experience working with SKOS in archaeological field projects.

Keywords:
SKOS, Ontologies, Thesauri

1. Entering Textual Information in Database Systems

There are several options for entering textual information in databases. Best-known are free-text fields, where uncontrolled strings may be entered. The benefit of this solution lies in the high flexibility of data entry, giving users the freedom to decide what information they enter on specific objects and how. This may be useful in specific cases, but disadvantages may outweigh advantages. Among other things, experience has shown that users tend to enter the same information in any number of different ways. One user will record a specific piece of pottery as “attic red figured pottery,” and the next will use the German phrase “attisch rotfigurige Keramik” or its abbreviation “RF.”, this will be supplemented by various data-entry errors (due to mistyping, etc.). That usually leads to many different labels for the same kind of artefact. Consequently, a reasonable analysis on this data is not possible, since you will never know whether you receive all the artefacts you were actually looking for. Hence a free-text field is not the best solution for information needed for use in further queries.

Controlled lists, such as those which are often used to provide default values for specific data-fields in Microsoft Access or FileMaker, offer a second solution for data entry. This method provides greater control over data and thus of the results of analysis, and in this sense may be preferable to free-text (Eckkrammer, Feldbacher, and Eckkrammer 2011: 82), but is also subject to some problems. Such lists tend to become so long that it is sometimes difficult to find the right term and, as a result, users often re-enter already existing terms either because they cannot find the one they are looking for, or else because they have a different opinion of what to name a specific type of artefact. So again: data analysis might not deliver the right result.

Another drawback is the list-form of the vocabularies. Generally it is not possible to map a complex multi-layer typology or classification onto these lists. Since it is also not possible to use different terms to describe the same thing, a decision must be made regarding the correct
naming. Multilingual vocabularies are also not possible.

Hierarchical thesauri have already been implemented in several database systems like Dyabola (http://www.dyabola.de/) or EH. But, since these were all monohierarchical and monolingual vocabularies, users were forced to decide for one term for every specific object. Also, use of a strict hierarchy requires consensus on the proper place for each term within the hierarchy.

If no consensus can be reached on naming and place, our experience has shown that the vocabulary will lack acceptance and users will use other free-text-fields to store information in ways not foreseen by the database designers.

Such problems may not seem very important in small datasets, but in the age of “big-data” (Anonymous 2008) they lead to useless data-piles without the possibility to properly analyse (or “mine”) the information they contain (cf. Baines and Brophy 2006: 236).

From such observations one might conclude that existing methods of entering textual data do not fulfil the requirements of modern, semantic-web-enabled database-systems.

2. The Simple Knowledge Organization System

As a solution for this emerging problem the W3C (http://www.w3.org/2004/02/skos/) developed the Simple Knowledge Organization System as a formal RDF/XML based language for representing thesauri, classifications, typologies and any other kind of controlled vocabularies. In the last years some semantic-web-oriented projects like STAR/STELLAR (May, Binding, and Tudhope 2011) have started to convert their vocabularies into SKOS to enable interoperability between different datasets.

SKOS offers a number of advantages regarding to the lists, free-text and monohierarchical thesauri. One of the main advantages of SKOS is the fact that it is concept-based. All of the connections between the thesaurus contents are defined by the concepts, thereby making it possible to use different terms and different languages to refer to the same concept. The concept consists of a URI which is in the best case a persistent identifier which enables the connection of the content of heterogeneous datasets.

The concepts serve as a kind of container which includes all additional information. The user interface generally shows the different preferred and alternative labels associated to the concept. The concepts can be connected by several semantic relations which are distinguished between two basic categories: hierarchical and associative. Hierarchical means a concept is more general than another. For example the concept “weapon” is more general than the concept “sword”. In SKOS, “weapon” is a broader term for “sword” and, reciprocally, “sword” is a narrower term for “weapon” (Fig. 1).

An associative related term indicates that two concepts are related. In this case one is not more general than the other. “Sword” is a related term not only to other bladed weapons like “knife” and “dagger,” but also to words like “Knight.” In addition to conceptual relations, the documentation properties are very important for the development of archaeological thesauri. Since every concept can be connected with different kinds of notes and examples which define what a specific concept actually means, if there is no consensus on where to place a specific concept in the thesaurus, then it can be connected to several different concepts.

This multilingual and multihierarchichal thesauri represent the reality both of archaeology and natural language (cf. Baines and Brophy 2006: 239-240) far better than the strict lists and monohierarchical and monolingual
vocabularies. Implementing SKOS-thesauri in database-system should lead to a greater acceptance of controlled vocabularies.

3. Implementation of SKOS-Thesauri in an Archaeological Database

In archaeology, SKOS-thesauri have so far only been used to link data in existing semantic-structures, as in the STAR/STELLAR-project. Our aim was to use SKOS in an environment for documenting archaeological fieldwork from scratch.

4. Tools for Generating SKOS-thesauri

The first problem we had to overcome was the lack of a proper tool for the development of new vocabularies. The SKOS Editor in Protégé (http://protegewiki.stanford.edu/wiki/SKOS_Editor) did not suit our purposes. The code is difficult to parse because it is not very well formed. The program also tends to make files unreadable. Most of the time, when Protégé crashed, it took hours to find the bug in the code which is not well formed and difficult to read. To counter this problem we wrote a script which brings the code into an appropriate form. But that was also not very intuitive to use, especially for archaeologists who are not used to command line tools and XML-Code (cf. Carver and Lang 2013). So the only proper solution was to develop a tool which forms proper code and was intuitively usable by people (i.e. not just IT-specialists).

5. The TuSaurus-Tool

We call our software TuSaurus (Fig. 2), in honour of the Technical University at Dortmund where the first version was developed as a project in a graduate course. TuSaurus was improved by the members of the ArcheoInf-group (http://www.archeoinf.de) and has now been used for several months by a number of archaeologists involved with the project.

The programming was done in Java using the Swing GUI widget toolkit. The TuSaurus tool fulfils SKOS-specifications and helps
archaeologists develop their own vocabularies quickly and properly in a standardised format without being confronted with XML-code or W3C-specifications (Fig. 3).

We are currently (April, 2012) revising the software to eliminate some bugs and increase stability. We will also add some export and import functions. English, Italian and Spanish-language versions of the interface will also be available in the next few weeks. Finally, TuSaurus will be made available as a fully documented open-source-software in the course of the next month.

6. The SKOS-thesauri in the ArchGate-Project

For documenting archaeological field projects we offer an integrated GIS-database-system named ArchGate (http://www.archgate.de) which consists exclusively of open-source-components. ArchGate is presently used by ten archaeological field projects in the Mediterranean area. The ability to use multilingual vocabularies greatly benefits such projects since work is usually conducted by international teams with partners from the country where the excavation or survey is conducted. Thus there is usually no common language for documentation, and in the worst cases this situation again leads to unconnected database entries in several languages for the same kind of object. As has been noted, SKOS offers the possibility of using multilingual vocabularies, with the result that all users are able to enter and query data in their own language (Fig. 4).

Below we will describe how we use the SKOS-thesauri for managing the content and the interface in the ArchGate-environment.

Because our database consists of relational PostgreSQL tables, it was not possible to implement the RDF-XML datasets directly. We therefore decided to parse the thesauri in SKOS-format into several tables linked by IDs. A tool offering some basic vocabulary entry functions has been built into the ArchGate interface, and there is always the possibility to import new thesaurus-trees (like existing typologies or classifications) into a thesaurus.
The interface only shows the labels (Fig. 2), and when starting the program the user has to decide which language he or she wants to use. In the database itself nearly all the textual content is controlled by the SKOS-vocabulary, thus allowing fast and consistent data entry (Fig. 5).

Due to the flexibility inherent in SKOS, this approach has been widely accepted by ArchGate users. Recording data in this standardised way increases the reliability of all query results by always containing all of the relevant information.

In addition to the content, the interface is also controlled by the SKOS-thesaurus. All the window, button, tab and box names are connected to SKOS-concepts. As a result, even the GUI can easily be switched to another language, and this approach makes it easy to customise the ArchGate interface in order to meet the requirements of new projects.

7. Conclusions and Experiences

The open and standardised format of SKOS makes it easy to adapt to specific requirements. by offering a wide range of semantic relations between the different concepts, the great flexibility when composing vocabulary makes it ideal for use in archaeology where (as has long been recognised; cf. Clarke 1978: xv) , in any many cases, there is little or no consensus on the proper classification or naming of a given object. Thus, by using a SKOS-thesaurus, we can begin to address the lack of a common language in our discipline.

Another benefit (one which was also shown by the STAR/STELLAR-project) is the easy export of the standardised vocabularies into semantic structures. Also the use of existing thesauri in your own data-structures is easy to handle. So there is no need to gather a new vocabulary for every new database as it has been so far. Unfortunately there is, as yet, no common place for storing thesauri with a common URI as a persistent identifier. At present, each project must develop its own vocabularies linked to its own URIs. Tackling this lack of a common repository for storing archaeological vocabularies with a persistent identifier for each concept will be one of the main issues of the SKOS-community in the future.

Beside the infrastructures, the tools also need to be improved. We hope that our TuSaurus-tool will help overcome the problem that – until now – some experience with XML-code has been necessary for creating SKOS-vocabularies, because so many archaeologists seem to have an innate fear of coming into contact with any pieces of code. Since the development of a SKOS-thesaurus will be much more time consuming than the conventional way of data entry using word-lists or free-text entries, we have to convince the archaeologists of SKOS's benefits by offering user-friendly tools and environments (Carver and Lang 2013).

References


Modelling Imperfect Time in Datasets

Koen Van Daele
Flanders Heritage Agency, Belgium

Abstract:
This paper takes a closer look at the problem of handling imperfect temporal information in historical databases. We inspect and define the nature of this problem, mainly originating from dealing with incomplete and vague information. Then we look for possible solutions in current scientific literature. We focus our attention on rough or fuzzy sets. We take a closer look at answering questions with these sets in a temporal context. We discuss how we implemented fuzzy time intervals in a PostgreSQL database. And finally, we discuss some test results that help us determine which method of answering fuzzy temporal questions is suitable in a live database in a production environment.

Keywords:
Temporal, Fuzzy Time Interval, Rough Time Interval, Database

1. Introduction

When constructing databases for applications dealing with history, archaeology or other related disciplines, we are often faced with incomplete or vague information. Our organisation maintains a large database of heritage in Flanders. This database contains records about archaeological findspots, historic buildings and landscapes. This information has an important temporal component. When trying to capture this in a database we are faced with several problems (Nagypál and Motik 2003).

First off all, we are faced with the problem of uncertainty. This problem manifests itself through information that we could know, but don’t. Among other things, the database allows heritage researchers to record a birth date and a death date for architects. Sometimes this information is only partially known. For instance, we know the year a person was born but not the day. We need a way to record this partial, but valuable, information.

If we take a further look at dates in a historical context we also come across some things that can’t be accurately dated because they have no clear beginning or end. A classic example of this is the industrial revolution. Nobody is able to pinpoint the exact day this started (or ended). Every researcher might have a different opinion about this. The same kind of problem arises with eg. archaeological cultures where we might be faced with a gradual appearance or decline. Therefore we call this type of temporal information subjective.

A special case of uncertainty that deserves mentioning is that of granularity. As humans, we talk about time in different scales. For some things we talk in terms of dates, for others in terms of hours, minutes or even seconds and sometimes we talk in terms of years, decades or centuries. If we eg. know that an event took place from 1940 to 1945 then this information is precise on a scale of years. But if we move to a scale of dates, it becomes imprecise. In effect, any temporal specification made in natural language will become vague if we refine the granularity of the temporal axis sufficiently.

For this paper we set out to find a way to represent uncertain birth and death dates in a database containing biographical information about architects. We do not only want to store this information (which could always easily be done in a text field), we also want to be able to query it. While we do focus on the problem of uncertainty, we do have other databases that deal with temporal information. For some of
those it is imperative that we can also represent subjective dates. So, we are looking for a solution that can represent both uncertainty and subjectivity.

2. Research

Our first option for handling imperfect temporal information is using a sharp set. In such a set there is a membership function $A: U \rightarrow \{0,1\}$. This membership function determines for a certain element $x$ whether or not it is part of the set, if $A(x) = 0$ then $x$ is not part of the set, if $A(x) = 1$ then $x$ is part of the set. A sharp set used to represent temporal information is known as a Sharp Time Interval (STI). We can use such an interval to represent uncertain time points, but not uncertain intervals. We can eg. represent the fact that a person was born somewhere in the year 1900 by stating that this person’s date of birth is $B = [01/01/1900, 31/12/1900]$, indicating that the birth took place somewhere in this year. If we try to do the same with this person’s lifespan, we lose a certain amount of information. Suppose we know a person was born in 1900 and died in 1995. We could encode this information as the interval $L = [01/01/1900, 31/12/1995]$. This interval would then either represent this person’s lifespan or his possible lifespan. However, we can not indicate that we know that this person might not yet have been alive in 01/05/1900.

To remedy this, we went looking into computer science literature for a possible solution. In doing so, we came across rough and fuzzy set theory (Qiang et al. 2010, Dubois and Prade 2000).

A rough set is a special kind of set where a distinction is made between a lower and an upper approximation. The lower approximation is the set of all elements that are certain to belong to the set. The upper approximation is the set of all elements that might belong to the set. The difference between these two sets forms a boundary region of elements for which we can not conclusively decide whether or not they are a member of the set. A rough set can be used to model uncertain information about a person’s life.

Suppose we have a person for whom we know he was born in 1900 and died in 1995. This leads us to a rough set ($L$) consisting of a lower approximation $L = [31/12/1900, 01/01/1995]$ that contains all the days we are certain he was alive. The upper approximation $L = [01/01/1900, 31/12/1995]$ contains the days he might have lived. There are also two boundary regions $B^-$ and $B^+$. These form the set of days that he might have been alive. The only difference between these two sets is that $B^-$ precedes $L$ and $B^+$ succeeds it. We call a rough set used to express temporal meaning (figure 1) a rough time interval (RTI).

A fuzzy set on the other hand, is a set where a membership function $A : U \rightarrow [0,1]$ indicates to what degree an element $x$ in $U$ belongs to the set. There are elements that certainly belong to the set and elements that certainly do not. There are also elements that might belong to the set. We can indicate the degree of their membership as a value between 0 and 1 (as opposed to a rough set, where we have elements that might belong to the set, but we can not express how much they belong). We call the set of elements that definitely belong to the set (elements $x$ for which $A(x) = 1$), the core (C). The set of elements that might belong to the set (elements $x$ for which $A(x) > 0$ and $A(x) < 1$), we call the support (S). Elements for which $A(x) = 0$ certainly do not belong to the set.
fuzzy set used to express temporal meaning (figure 2) is also known as a fuzzy time interval (FTI). With a FTI we call the part of S that is before C the Fuzzy Beginning (FB). The part of S that is after C we call the Fuzzy End (FE). Because an FTI allows us to express a degree of membership, it is inherently more suited to represent subjective information. For example, we can represent the slow evolution that was the industrial revolution. Therefore, we choose to use FTI’s for the remainder of our work.

Of course, being able to represent uncertain and subjective temporal information is one thing. We would also like to query this information. Allen (Allen 1983) determined a set of 13 possible relations between two sharp time intervals (STI). Of these 13 relations (table 1), 12 can be divided in 2 groups that are each other’s inverse. For if A is before B, then B is after A. Or if A overlaps B, then B is overlapped by A. The last relation, equals, is it’s own inverse. For if A equals B, then naturally B equals A as well.

For our purpose we have added 5 extra composite relations, because the 13 Allen relations are too detailed for most heritage researchers. Eg. we have added a relation bef that is a combination of before and meets. So if A is bef B, this means that A is either before B or it meets B. We did this because we found that heritage researchers considered it odd that eg. the Interbellum would not be before World War II according to the Allen definitions. Similar to the regular Allen relations, we also added an aft relation as the inverse of the bef relation. We also added the “intersects” relation that was suggested in (Nagypál and Motik 2003). This relation determines if two FTI’s have any point in time in common. In essence it is the same as (overlaps(A, B) OR during(A, B) OR starts(A, B) OR finishes (A, B) OR overlapped by(A, B) OR contains(A, B) OR started by(A, B) OR finished by(A, B) OR equals(A, B)). This relation, as with the equals relation, is also it’s own inverse. For if A intersects B, then B also intersects A. We find that this is the relation researchers are most interested in in a large dataset. It is this relation that can be used to answer questions like: “Give me all architects that were alive in some part of the 19th century”. In the end, this left us with 18 relations that we wanted to implement for FTI’s.

Although we did an extensive survey on literature on defining these Allen relations for RTI’s, we could not find any information on this topic. This would certainly be an interesting area for further research.

We did come across several different algorithms for determining fuzzy temporal relations between FTI’s. Each of these methods specifies the relations between two fuzzy temporal relations as a fuzzy relation. This means that the result is again expressed as a relation.

<table>
<thead>
<tr>
<th>Relation</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>before</td>
<td>( b(A, B) = a^+ &lt; b^- )</td>
</tr>
<tr>
<td>overlaps</td>
<td>( o(A, B) = a^- &lt; b^- \land b^- &lt; a^+ &lt; b^+ )</td>
</tr>
<tr>
<td>during</td>
<td>( d(A, B) = a^- &gt; b^- \land a^+ &lt; b^+ )</td>
</tr>
<tr>
<td>meets</td>
<td>( m(A, B) = a^- = b^- )</td>
</tr>
<tr>
<td>starts</td>
<td>( s(A, B) = a^- = b^- \land a^+ &lt; b^+ )</td>
</tr>
<tr>
<td>finishes</td>
<td>( f(A, B) = a^- = b^- \land a^+ &lt; b^+ )</td>
</tr>
<tr>
<td>equals</td>
<td>( e(A, B) = a^- = b^- \land a^+ = b^+ )</td>
</tr>
<tr>
<td>after</td>
<td>( a(A, B) = b(A, B) )</td>
</tr>
<tr>
<td>overlapped by</td>
<td>( o(A, B) = o(B, A) )</td>
</tr>
<tr>
<td>contains</td>
<td>( c(A, B) = d(B, A) )</td>
</tr>
<tr>
<td>started-by</td>
<td>( s(A, B) = s(B, A) )</td>
</tr>
<tr>
<td>finished-by</td>
<td>( f(A, B) = f(B, A) )</td>
</tr>
<tr>
<td>bef</td>
<td>( bef(A, B) = a^- \leq b^- )</td>
</tr>
<tr>
<td>dur</td>
<td>( dura(B, A) = a^- \geq b^- \land a^+ \leq b^+ )</td>
</tr>
<tr>
<td>intersects</td>
<td>( inters(B, A) = a^- &gt; b^- \land a^+ &lt; b^+ )</td>
</tr>
<tr>
<td>aft</td>
<td>( con(A, B) = bef(B, A) )</td>
</tr>
<tr>
<td>con</td>
<td>( con(A, B) = dura(B, A) )</td>
</tr>
</tbody>
</table>

Table 1. Possible relations between 2 Sharp Time Intervals. \( A = [a^-, a^+] \) and \( B = [b^-, b^+] \).
number between 0 and 1, signifying to what degree this relation is true, with 0 meaning absolutely false and 1 absolutely true.

Nagypál and Motik (Nagypál and Motik 2003) were the first to propose such a method. Their method is very simple, and works rather well. However, it cannot be used for fuzzy temporal reasoning. When dealing with two STI's intervals, the Allen relation equals is reflexive (if A equals B, then B equals A). So, we would expect that a FTI interval is equal to itself: equals(A,A) = 1. However, using Nagypál and Motik’s definitions, we arrive at equals(A,A) = 0.5. Because of this we can not use their method for temporal reasoning. Eg. if we knew that before(A,B) = 1 and before(B,C) = 1, we still can not infer that before(A,C) = 1.

Schockaert, De Cock and Kerre (Schockaert, De Cock and Kerre 2008) adress a different way. They define a relation that expresses the extent to which a is long before b. They also define a similar relation that expresses to what extent a is before or at the same time as b. Using these two fuzzy relations they define fuzzy Allen relations. These definitions do appear to be very computationally expensive. They do offer the advantage of being useable for temporal reasoning. So if we know that before(A,B) = 1 and before(B,C) = 1 then we can actually conclude that before(A,C) = 1.

In (Schockaert, De Cock and Kerre 2006) they suggest an alternative that is computationally less expensive but only useable for trapezium shaped FTI’s. These FTI’s consist of only four dates; these mark the start and end of the support and the core. A fifth variable that can be taken into account in this algorithm is Lambda, being the maximum value the membership function F(x) can have for this interval. As with their other method, the possibility of temporal reasoning remains. The constraint on the type of FTI that can be used seems limiting at first, but it is our experience that these trapezoidal FTI’s can be used to represent most situations we regularly deal with in our databases.

3. Implementation

Having researched these possible solutions, we decided to try and implement the methods proposed by Nagypál and Motik and Schockaert. We wrote methods to encode an FTI as a LINESTRING object in a PostgreSQL database using the PostGIS extension. PostGIS was used because it offers operations like intersect, union and difference on linestrings. The resulting code is available under an open source license (https://github.com/koenedaele/pgFTI).

Every fuzzy relation was implemented twice. Once using the Nagypál and Motik method (NM), once using the alternative method proposed by Schockaert (S2). The Allen before method was also implemented using the more expensive Schockaert method (S1). Each of these implementations contains some guard clauses that shortcircuit the expensive calculations when the fuzzy relation is obviously true or false. Eg. when evaluating the Allen before relation between 2 FTI’s A and B, we could skip the actual algorithm if the support of A was complety before the support of B and just conclude that the relation was true.

To test and compare the three methods, we constructed a database containing biographical...
data for 716 persons. To represent each person’s life, an FTI was constructed. In doing so we also came upon a problem of missing data. If a person had a recorded birth date of 15/04/1804 we could safely assume that they are not longer alive. But for people born in the 20th century, we had to be a lot more careful. Based on our data we concluded that if a person was present in our database they must have lived at least to 20 years of age (since they had to have had the time to contribute in a significant way as an architect). We also decided that the average person in our database would not have lived to be more than a hundred years of age. In the end this led us to constructing FTI's shaped like figure 3. For this person we know that he died 10/12/1394. We assume hat he was at least 20 and possibly 100 at that time. So we end up with a FTI with a vague beginning between 10/12/1294 and 10/12/1374.

In our tests we queried this table of architects with each algorithm. We varied the time-intervals with which we queried to be able to observe the effects of the distribution of temporal information on the query results. Eg. most of our architects lived in the 19th and 20th century, none were born before 1.100 AD and none were born after 2.000 AD. Table 2 offers the results for calculating the before relation using NM, S1 and S2 using four different time intervals. From these results we learn that each method produces similar results, but they do differ tremendously in the amount of time it took to calculate this. The S1 method is clearly several orders of magnitude slower than the other two methods. The NM and S2 methods are more evenly matched. We can however observe that S2 seems to scale better when the number of people that has to be examined by the actual algorithm increases. This demonstrates that the guard clauses we build in to prevent the expensive algorithm running unnecessarily make a huge difference.

We also tested the other Allen relations for this dataset by comparing the NM and S2 implementations. To reduce the probability of unexpected fluctuations (eg. network latency), each test was ran 10 times and the results of these tests were averaged (Van Daele 2010). Table 3 shows the results for calculating all Allen relations for the date “01/01/1900” and our dataset. These further tests confirm that both algorithms produce similar, but not exactly alike, results. We can also clearly observe that the S2 method scales better. When the number of records for which the algorithm actually needs to run increases, the S2 method suffers

\[\text{Table 2. Comparing the fuzzy Allen before relation for NM(Nagypal and Motik), S1(Schockaert), S2 (Efficient Schockaert).}\]

<table>
<thead>
<tr>
<th>Date</th>
<th>NM</th>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/1001</td>
<td>0</td>
<td>0.12</td>
<td>0</td>
</tr>
<tr>
<td>01/01/1701</td>
<td>7</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>01/01/1901</td>
<td>132</td>
<td>0.58</td>
<td>419</td>
</tr>
<tr>
<td>01/01/2101</td>
<td>716</td>
<td>0.06</td>
<td>0.1</td>
</tr>
</tbody>
</table>

\[\text{Table 3. Comparing the fuzzy Allen relations for NM(Nagypal and Motik) and S2 (Efficient Schockaert) for the date of 01/01/1900.}\]

<table>
<thead>
<tr>
<th>Allen relation</th>
<th>NM</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>132</td>
<td>132 0.093</td>
</tr>
<tr>
<td>Meets</td>
<td>8</td>
<td>8 0.105</td>
</tr>
<tr>
<td>Overlaps</td>
<td>8</td>
<td>8 0.153</td>
</tr>
<tr>
<td>Starts</td>
<td>0</td>
<td>0 0.127</td>
</tr>
<tr>
<td>During</td>
<td>0</td>
<td>0 0.095</td>
</tr>
<tr>
<td>Finishes</td>
<td>8</td>
<td>0 0.100</td>
</tr>
<tr>
<td>Equals</td>
<td>0</td>
<td>0 0.096</td>
</tr>
<tr>
<td>After</td>
<td>207</td>
<td>207 0.100</td>
</tr>
<tr>
<td>Met by</td>
<td>0</td>
<td>0 0.070</td>
</tr>
<tr>
<td>Overlapped by</td>
<td>3</td>
<td>3 0.098</td>
</tr>
<tr>
<td>Started by</td>
<td>3</td>
<td>3 0.096</td>
</tr>
<tr>
<td>Contains</td>
<td>388</td>
<td>388 0.122</td>
</tr>
<tr>
<td>Finished by</td>
<td>0</td>
<td>8 0.136</td>
</tr>
</tbody>
</table>

\[\text{All tests were carried out on a remote PostgreSQL 8.4 database running on a Sun Solaris server (Sparc CPU) through a vpn tunnel.}\]
much less when it comes to performance than the NM method.

Our implementations for NM and S2 were all done using PostGIS because NM needs some spatial operations like intersect. S2 however does not need this. To make sure that both algorithms were operating in the same technical environment we implemented both of the algorithms using PostGIS. S2 could quite easily be implemented without relying on PostGIS, making this method even faster.

4. Conclusions

We were looking for a solution to handle uncertainty and subjectivity in temporal information in our database. To this end we looked at Rough Temporal Intervals and Fuzzy Temporal Intervals. Although both can handle uncertain temporal information, only the Fuzzy Temporal Intervals can cope with subjective temporal information (e.g. the gradual beginning of an archaeological culture). When it comes to determining the relations between temporal intervals, we rely on the relations defined by Allen in 1983. Some research has been done in determining these relations for Fuzzy Time Intervals. For Rough Time Intervals, this research is lagging behind. We implemented and tested several of the possible algorithms for determining the Allen relations between two Fuzzy Time Intervals. They provided similar results, but differ in the time required to compute these results.

In the end, which algorithm is best used, depends on the situation. If we are dealing with a large dataset in which the imperfect temporal information can be represented by a trapezoidal Fuzzy Time Interval, we want optimal performance and the option to do temporal reasoning, the alternative Schockaert (S2) method should be used. If performance is not an issue, trapezoidal Fuzzy Time Intervals are not an option and temporal reasoning is important, we should use the standard Schockaert method. Finally, the Nagypál and Motik method offers decent performance and can handle more complex Fuzzy Time Intervals, but it can not be used for temporal reasoning.

Acknowledgements

This paper is based on the research for my masterpaper at Ghent University (Van Daele 2010). I would like to thank my promotors, Guy De Tré and Antoon Bronselaer, for their advice and support.

References


**Distribution Analysis of Bone Remains in the Prehistoric Site of Mondeval De Sora (Belluno - Italy): Issues and Proposals**

Maria Chiara Turrini, Federica Fontana, Antonio Guerreschi and Ursula Thun Hohenstein  
*Università degli Studi di Ferrara, Italy*

**Abstract:**
This paper illustrates the methodology applied to the analysis of the spatial distribution of bone remains from a layer of Mondeval de Sora (Belluno - Italy). This is an exceptionally rich and well preserved high mountain site (2150 a.s.l.), with early Holocene deposits located in the eastern sector of the Italian Alps. The methodology developed has allowed to homogenize the number of remains collected by excavating the chosen layer with sub-grids of different sizes (10 cm, 20 cm and 33 cm) over the years. In addition it has been possible to assess on a quality level the impact of the error that is introduced when the construction of an isopleth map is preferred to a more accurate choropleth map.

**Keywords:**
Distribution Analysis, Choropleth Maps, Isopleth Maps, Mondeval de Sora

1. **Introduction**

Among the different types of analysis introduced in archaeology by the use of GIS the spatial distribution of finds has become very popular in the last decades. The specific aim of this approach is to highlight the existence of possible distribution patterns which may reflect the spatial organization of activities within human settlements. This aspect is particularly important when investigating prehistoric sites, especially of Palaeolithic and Mesolithic age, which only rarely yield dwelling structures. Therefore the reconstruction of settlement dynamics strongly depends upon the analysis and interpretation of the spatial relationships which interconnect different categories of archaeological findings (i.e. lithic artefacts, faunal remains, archaeobotanical remains, etc.).

The most immediate way to perform this analysis is to show the points that identify the remains in their correct relative positions within the excavated area. To do this we need to determine the coordinates (x, y and z) of each object. These coordinates will be used in the GIS for creating the map of the analyzed layer. Today this process is facilitated by the use of total stations that can assign Cartesian coordinates to each remain with a very simple measuring.

Nonetheless this operation is not always possible especially when the finds are very abundant and it becomes too time consuming. In these cases the use of a grid and a sub-grid reference system is preferred. As a consequence in order to perform spatial distribution analysis it becomes necessary to adopt another method. Particularly we must transform the analysis of the spatial distribution of points i.e. analysis of point patterns into an analysis of area data i.e. analysis of the number of objects found within a unit area (Bailey and Gatrell, 1995).

The visualization of area data requires a classification of the considered grid system based on the number of objects found in each area, thus making it necessary to solve the problem of the choice of the class limits (Connolly and Lake, 2006). A shortcut is made available by the ease of use of some calculation modules implemented in many vectorial GIS and is given by the method of interpolation of point values that can be performed after having paired the centroid of each elementary area.
with the number of objects that it contains. In this way an isopleth map of the distribution of the finds is created (Rinaldi, 2010). Although this system appears very rapid it results in a spatially continuous data analysis (Bailey and Gatrell, 1995) instead of an area data analysis.

Therefore when this type of analysis is adopted it becomes crucial to understand and quantify the error that is introduced by its use. In order to investigate this aspect Authors have applied this methodology to a layer of the site of Mondeval de Sora (Belluno - Italy).

2. The Site of Mondeval de Sora (Belluno – Italy)

Mondeval de Sora lies on a terrace in the upper Fiorentina valley (Belluno, Dolomites) at an altitude of 2150 m a.s.l. This site represents the richest and best preserved high altitude early Holocene deposit of the eastern side of the Italian Alps (Alciati et al., 1992; Fontana and Guerreschi, 1998; Fontana et al., 2009; Rinaldi G., 2010) (Fig. 1). It consists of two sectors (I and III), which are located under two different overhangs of a large boulder of dolomite (Fig. 2). In particular sector I, which was explored over an area of approximately 60sqm, has revealed a complex stratigraphic sequence about 50cm thick which extends - with various lacunae - from Early Mesolithic to the sub-present age.

The Early Mesolithic (Sauveterrian) evidence contained a paved area made of local tufa slabs (SU 14), delimited by an arrangement of blocks of dolomite stones (SU 33) and a sub-circular structure, interpreted as a hearth (SU 32). Two anthropic layers covered these structures being located respectively in the inner (SU 8) and in the external part of the shelter (SU 31). Stratigraphic Unit 8, which is the object of this analysis, was constituted by a 20cm deep dark brown-black silty-sandy layer, rich in lithic artefacts, faunal remains and charcoals. It has been attributed to the middle and recent phases of the Sauveterrian on the basis of the techno-typological features of the lithic assemblage and of a radiocarbon date obtained from a charcoal sample (GX-21788-9185 ± 240 BP, 9175-7731 cal BC, 2σ) (Fontana and Vullo, 2000). Results from previous studies have allowed Stratigraphic Unit 8 to be interpreted as a palimpsest due to several occupation phases. The presence of dwelling structures indicates the seasonal residential character of the site which - according to its location - was occupied by hunter-gatherers during the spring and summer seasons, while results of the analysis of the archaeological record delineate a functional emphasis towards the procurement and exploitation of animal carcasses.
For the purpose of this work the spatial distribution of faunal remains coming from the Layer 8 has been analyzed. The analysis reported in this work has the purpose to compare the use of different methodologies and it is preliminary to the interpretation of the spatial distribution maps obtained. The final goal of our work is to compare results of spatial analysis of faunal remains with those of lithic artefacts, which was the object of a previous work (Vullo et al., 1999), in order to detect use patterns of domestic space with the site. This will be the object of following publications.

3. Methodology of Data Processing

3.1 Homogenization of Elementary Areas

The main problem connected to the spatial analysis of SU 8 at Mondeval de Sora was that this layer was excavated by sub-grids of different sizes within the main grid of 1m. At first a sub-grid of 10x10cm was used in order to have a higher detail of the spatial distribution of the finds. However since this method revealed to be too time-consuming the decision was taken to excavate by a sub-grid system of 20x20cm and then 33x33cm. This last sub-grid has also been used when the density of the finds was lower. In addition where the layer was thicker it has been excavated by artificial levels (up to three) not always applying the same mesh size. The result was a high discrepancy both horizontally and vertically of the sub-grid reference system.

Therefore the first necessary step of the work has consisted in the conformation of the different sub-grids to a homogeneous grid of 33x33cm, assuming that the finds were evenly distributed within each sub-square. Figure 3 shows that each sub-square of size 33cm numbered from 1 to 9 includes a non-integer number of sub-squares of size 10cm and 20cm. By applying the topological overlay tool of the GIS the percentage area of each sub-square of smaller size contained in each 33cm size sub-square has been calculated.

As an example Table 1 gives a reference for sub-square 1/9. SU 8 is thus composed by 207 sub-squares of size 33cm. This quantity allows a good interpolation to be performed.
3.2 Database

A database with the required information for the analysis has been then realized starting from the database created for the study of bone remains. This was possible since the latter is a highly structured database. Among other information each remain is related to the mesh of size 1m and the relative sub-square.

The first operation consisted in adding up all the bone fragments belonging to each sub-square of size 33cm, 20cm and 10cm in order to quantify the total number of finds for each sub-square. The values obtained were then summed up in percentage as in Table 1 in order to obtain the total number of finds for each sub-square of side 33cm. The standardization of the finds distribution with respect to a sub-grid of side 33cm has thus been obtained. Considering a total of 207 sub-squares only 47 (22.22%) did not contain any bone fragment.

Finally the coordinates of the centroid of each sub-square with 33cm side, which were previously identified by applying one of the GIS tools, have been entered in the database.

Table 2 shows the fields of the database which has thus been created. The label refers to the identification code of the sub-square of side 33cm (* / 9) inside the mesh of 1m (n).

3.3 Sub-squares Classification

As a consequence of the excavation strategy adopted and the resulting structure of the database the data obtained are area data. Therefore the variation of the number of findings that could originally be considered as continuous has thus been discretized according to a regular grid of size 33cm. In this grid system each sub-square is characterized by a defined value, i.e. the number of bones found inside. Thus it was considered that the mere display of the grid in which the sub-squares are coloured according to their value may be a good way to visualize the variation of bone fragments concentrations.

A previous work by Vullo et al. (1999) aimed at analyzing the spatial distribution of the lithic industry from SU 8 with a raster GIS has solved the problem of the homogenization of the elementary areas by applying a method which is similar to the one described in this work. The distribution maps of the various findings divided by technological categories elaborated by these Authors are choropleth maps. However the criteria used for the choice of the class limits have not been specified by Authors which let us assume that the choice has been made subjectively and based on their personal experience. In this work the system of the coloured sub-squares has also been applied in order to visualize the variation of the objects concentration.

In order to apply this system a classification of the sub-squares has been carried out using the query tool. The problem addressed is that of the univariate numerical classification where the considered variable is the total number of bone fragments for each sub-square.

The resolution of this problem envisages the clustering of the sub-squares on the basis of the class limits obtained by studying the data structure. There are several ways to define the limits of the classes to be adopted based on the data distribution function (Connolly and Lake, 2006).

In this case the data obtained have a positive skewed distribution for which Connoly and Lake (2006) suggest the use of the geometric intervals method. Nonetheless this method allows the low values classes to be seen in detail while it combines those with higher values. As a consequence it hides the information which is considered as the most interesting for this case study.

<table>
<thead>
<tr>
<th>Label n-* / 9</th>
<th>X Coordinate</th>
<th>Y Coordinate</th>
<th>Number of bones</th>
</tr>
</thead>
</table>

Table 2. Fields of the new database.
Therefore we decided to apply the quantile method. By splitting the series of data in equal parts it allows a uniformly detailed view. In particular the deciles were used. As a consequence of the fact that the sub-squares with a zero value are around 20% the unification of the ninth and the tenth classes has occurred.

Figure 4 shows the result of the classification of the sub-squares obtained with a tool of the GIS (choropleth map) and the related class limits.

3.4 Application of Kriging

After the creation of the choropleth map, which is considered as the closest to the real distribution of the finds based on the data processed as described above, the alternative method of assigning the area numerical value to the centre of gravity of each sub-square has been applied. The interpolation between all contiguous points has then been performed using the algorithm of Krige (Connolly and Lake, 2006) in order to obtain the isopleth map of the distribution of the findings. The result obtained is shown in Figure 5 where the values of the isopleths are those that the software gives by default without any further processing by the operator.

By comparing Figure 4 and Figure 5 we can notice that in the second one the sub-squares which contain the highest amounts of bone fragments are better highlighted thus allowing a more immediate visualisation of the fragments concentrations.

Of course the visualization changes according to the class limits chosen. In order to highlight the correctness of the graphical result obtained by the method of kriging the layout of the isopleth map was created by using the class limits obtained with the method of deciles. The map thus obtained (Fig. 6) is substantially comparable to the choropleth map (Fig. 4).
4. Discussion and Conclusions

In spite of the very similar graphical results of Figures 4 and 6 we point out that the method of interpolation is conceptually incorrect. Indeed if we consider the strategy adopted for data collection i.e. the number of finds for each sub-square, the final processed data are area data. In particular, it should be noted that between the value of a sub-square and that of its contiguous ones continuity does not exist but the values are discretized.

On the contrary when the method of interpolation is applied the starting condition is that the variable analyzed is continuous all over the investigated area, i.e. in the whole area occupied by SU 8. Virtually the introduced error consists in treating as a discrete variable whose value is valid for an entire area (the sub-square) as a continuous variable “measured” at a precise point (the centroid of the sub-square).

In order to analyze the distribution of the finds we should be able to identify the position of each find in the investigated area. When coordinates cannot be assigned it becomes necessary to process the area data obtained by building the choropleth map of the investigated area.

In the case study which has been presented here the choropleth map built (Fig. 4) has been used as a comparing map in order to evaluate the error that is introduced when an isopleth map is built instead of the choropleth one. When creating the isopleth map it is assumed that the number of finds for each sub-square is assigned to a point internal to the sub-square (in this case its centre of gravity) in order to allow data processing as if we would deal with a continuous variable.

The result obtained is conceptually wrong since between the isopleths passing through the centroids of two adjacent sub-squares there exist other isopleths with intermediate values. This is obviously in contrast with real data.

Nonetheless if we do not consider the isopleth map as a modelling of data but as a means for data visualization the result can be accepted.

In this paper Authors demonstrate that the isopleth map may be graphically very similar to the more correct choropleth map when the class limits are conveniently chosen.

The facility of application of the interpolation process often leads operators to use this tool since the classes obtained by default usually give the best visualization of the distribution of archaeological objects (Fig. 5) and do not require any reasoning concerning the choice of class limits. This choice is very important in order to avoid a false rendering of the remains distribution.

References


Places, People, Events and Stuff; Building Blocks for Archaeological Information Systems

Paul J. Cripps
Wessex Archaeology, UK

Abstract:
Archaeological information is by its very nature complex and uncertain. Typically, databases (when used) are used to record a ‘perfect’ and simplified version of the available archaeological information; there is little room for multivocality, uncertainty is reduced to a value qualifier and fundamental concepts are semantically indistinct. It is time for archaeological information systems to move forward with respect to core concepts of subjectivity, multivocality, temporality and uncertainty. This paper will examine how these concepts can successfully be modelled and made explicit within archaeological information systems, such theoretical constructs being independent of any modelling language or implementation platform. This paper will then present related work on a Spatial Data Infrastructure (SDI) being developed by a commercial archaeological unit based around these core concepts. Finally the paper will explore avenues for improving exchange of information, specifically ways in which SDI principals and technologies could be applied to UK local and national heritage records.

Keywords:
Spatial Data Infrastructure, Ontology, Semantics, Archaeological Information System, Data Management

1. The Nature of Archaeological Information

As archaeologists, our understanding of the past is based on the information we gather through our activities including excavation, survey, analysis of finds and objects collected and documentary research. This is then combined with layers of interpretation to produce narratives regarding the past. Traditional approaches to the creation of such narratives typically involve moving from initial observations and assessments to more detailed analysis and finally interpretation with a variety of documentation produced along the way. The nature of any information gathered and associated documentation is varied to say the least.

Taking archaeological information gathered through the process of excavation as an example, it is possible to see a gradual shift in methods and outputs. Historical excavations often resulted in very little structured documentation with the final publication by the site director being the definitive record. More recent excavations tend to produce more detailed records, influenced by the idea that excavation is a destructive process and there needs to be some form of preservation by record. More detailed, structured records have also been encouraged through the formalisation of excavation, post-excavation and related activities through management protocols such as MAP2 (English Heritage, 1991) and MoRPHE (English Heritage, 2006), a response to assist with the scale of modern archaeological projects and programmes, often involving complex team structures and spanning multiple organisations over many years.

In short, there has been a transition away from an almost complete absence of structured primary records with a single ensuing complete, authoritative narrative. The direction of travel has been towards more detailed, structured records (ie what is often referred to as a ‘primary’ archive dataset) often associated with metadata to describe and support any data. This ‘primary’ data is then used to produce derived products and develop narratives to convey understandings regarding the archaeological record.
The concept of primary data is one borrowed from the hard sciences where processes of experimentation and observation result in empirical, repeatable data points. These data points will support or undermine particular theoretical views in a consistent manner; the relationship between primary data, hypothesis and eventual theory is robust, well understood and explicit. Such data points typically lend themselves to structured forms of data management and analysis particularly within information systems. Unfortunately, the treatment of archaeological information in this way can be problematic and this has been the subject of extensive discourse within the post-processual movement of the past few decades; the application of reductionist, particularly statistical, overtly scientific approaches to archaeological data does not adequately model the richness of human existence and can lead to overtly functionalist theories regarding the past. There can be many reasons behind archaeological observations; it is not the case that there are universally applicable rules, laws or equations governing the spatio-temporal patterns we, as archaeologists, observe and record. Furthermore, the production of archaeological observations is not repeatable, archaeology being a destructive process, hence it is impossible to talk about archaeological records being empirical in the scientific sense. The construction of archaeological records is more akin to the classificatory and descriptive activities found in sciences such as ecology, paleontology and much of biology. Archaeological information is now therefore broadly understood to be partial and imperfect, often internally contradictory.

1.1 Assertion, Inference and Multi-vocality

A significant aspect of archaeological narrative is that it is possible to construct multiple narratives from the fragmentary remains archaeologists have to work with. It may be that stratigraphic sequences are broken and incomplete, physical relationships between excavated features may be unclear and ambiguous, documentary records may conflict with excavation records. Furthermore, spot dates based on pottery sequences may not correspond with other available evidence and this may be due to a number of reasons including but not limited to problems with a particular documented pottery sequence or issues relating to pottery identification and classification or discrepancies elsewhere concerning other aspects of the available knowledge base.

The scenarios outlined above make archaeological information rather more complex than a simple collection of observations or empirical data points. Any piece of archaeological information will have been selectively constructed by an archaeologist within some social context. The available evidence may well support multiple narratives and this is not a problem per se, it becomes a problem only when we attempt to work with archaeological data in ways which do not adequately support this multi-vocality. Crucially, this is of prime importance when we come to undertake the process of inference and reconstruction of a perception of the higher order structures from the remains observed in the field; in other words as we move from detailed records concerned with individual deposits through the processes of grouping and phasing to infer sequences, structures, sites, monuments, complexes and entire landscapes.

If archaeological data is stored in a typical flat file or relational database structure, it may not be possible to record the necessary richness; this is particularly so where rigid data structures enforce single values for data objects. For example, if we take a context based set of excavation records entered into a database, it may be a restriction that a recorded context may only be assigned to a single group object which in turn belongs to a single phase of activity. In such a scenario, it is not possible to record multiple assertions as to how the recorded archaeology fits together, only a single world view is possible. The data model is in effect rather hierarchical and monolithic, there
being no potential for recording multiple views of the data with the information system.

One solution to this has been to change the way in which the data structure is modelled so as to implement different forms of relationship within a relational structure, in effect to broaden membership of groups to allow for a more descriptive database. In such systems it becomes possible to assign multiple overlapping memberships and there may even be audit trails to indicate who created records and when. The problem then arises as to which sets of records to use, which are the ‘correct’ records, the most ‘up-to-date’ records, the implication being that there can still only be one ‘true’ record. Without a data model capable of adequately describing not only the archaeological data contained within records but the archaeological processes used to generate records, any database is limited in terms of its capabilities and suitability for reporting, assessment and analysis and ultimately its suitability as an archive.

1.2 Events and Archaeological Processes

Drawing heavily on Object Oriented programming techniques and Object Relational data models, the notion of event based modelling can also take into account the descriptive qualities of ontological models such as the CIDOC CRM (Crofts, 2003). Of course, there are many, sometimes complimentary, sometimes contradictory, approaches to event based modelling and as a generic concept independent of any particular technology, event based modelling can be seen as a powerful tool for archaeological information systems development. If we do then leverage a common ontological framework to support this, for example the CIDOC CRM, we can also facilitate enhanced interoperability between heterogeneous archaeological datasets. Our Event based model(s) become semantically clear and explicit rather than individualised and disjointed and we can relate groups of archaeological processes to one another more effectively.

In general then, if we define Events as explicit data objects within archaeological information systems, we can then define any other data object(s) as being the product of some form of Event. We can model not only the data objects (ie the archaeological records) themselves but any events associated with them and any other data objects which participated in those events. In other words, we can record not only the data we ordinarily would but the sequences of (usually implicit) events in the past as well as inference chains in the present leading to the production of the data. We do need to be careful about the level of granularity and detail we wish to record and pragmatic decisions will determine what is recorded, with any implementations needing to provide enough flexibility to facilitate this; this can be problematic and the necessity to create superfluous records must be avoided, a lesson learned from archaeological systems which have tried to implement such a basis. Overall though, the use of an event based model has a number of benefits.

Firstly, our data models become far richer and more useful but without the need for extensive development of overly specific types of record. By not simply reducing information to implicit, often poorly defined properties or attributes of arbitrarily defined entities, generally replicating historic paper based recording systems, the nature of the records becomes more explicit. Yet at the same time, there is no need to overly extend a complex data structure to account for myriad eventualities. We can record more detail, with simpler data structures and with improved semantic clarity. If we take a pot as an example, we can talk about the event of producing it and we can talk about the deposition event which led to it being where it was discovered. We can then talk about the excavation event in which it was discovered, classification events which assigned particular typologies based on specific type libraries and any conservation events it was then subjected to. Importantly, this does not require specialisation of an entity
relationship model to the n-th degree, this richness can be recorded quite simply using a minimal number of types of data object, notably a physical object and multiple events. Through inheritance of properties, and based on the CIDOC CRM concept of an event, any event will involve people or organisations and be spatially and temporally bounded, an ideal descriptive framework for not only events in the past (ie in archaeological/historical time) but events in the present (ie archaeologists doing archaeology); so we can know who classified what, when, and based on which evidence.

Secondly, by making our information systems simpler yet more explicit with greater semantic clarity, we facilitate novel uses of our data. In addition to knowing more about the construction of knowledge, the possibility of being able to then take advantage of automated processing and reasoning tools is facilitated (eg Vlachidis et al. 2010; Binding, 2010). Representing data in the form of RDF triples which can feed graphing tools such as jnet (Ryan, 2001) to dynamically view and amend stratigraphic sequences for example. Using Linked Data (Berners-Lee, 2006) approaches so as to be able to join up typological databases with site databases so that as interpretations change through time, it is possible to easily identify, inform and propagate change, facilitating re-evaluation of archival material based on current thinking. And all of the above forming components of integrating and aggregating distributed information systems imbued with semantic clarity and interoperability.

2. Archaeological Information Systems and Implementations

Whilst the theoretical implications of the approaches described above must be described and discussed, it is through implementation and use that benefits will be realised. Implementation has unfortunately been a major problem for archaeological information systems, with numerous examples of poorly modelled, poorly implemented databases which have hindered rather than helped research and reinforced the view that there is a fundamental problem with the use of digital, scientific techniques. Such techniques are often described using terms such as ‘reductionist’ and whilst this is a fair criticism in many ways, that is not to say that all such implementations are fundamentally and irreconcilably flawed. This is due, in part, to the way in which archaeological information systems have all too often been underspecified and then crudely implemented in many cases by archaeologists with little or no formal training in or knowledge about effective data management, the importance of the underlying data model and the resultant use, manipulation and presentation of data to suit the needs of assessment, analysis and reporting processes. The information systems have simply been used as digital versions of paper based archives and thus inherited all the limitations thereof but less of the restrictions, leading people to think they can do more because the data is in a digital system.

One view has been that it is not possible to dictate how an individual archaeologist should structure their data and indeed the CIDOC CRM, from where many of the concepts described previously in this paper have their origins, states that it is a descriptive ontological model aimed at being applied to existing data structures rather than acting as the basis for modelling itself. I would argue though that the approach being described here is not prescriptive nor is it restrictive. By specifying a core set of classes of objects which can be used to model and build an almost limitless number of data structures, it is hoped that a broad consensus can be reached that such an approach are indeed widely applicable and useful. Furthermore, such an approach can provide a common basic framework through which effective systems can be developed more easily by building on this robust core set of object classes.
If we now turn to the specifics of the approach. This approach was first proposed in 2004 (Cripps et al, 2004; Cripps & May; 2010) as a theoretical construct and some of the concepts date further back, for example the essence of the Framework Archaeology system (Andrews, et al. 2000). This paper aims to summarise and elucidate further through examples and experiences associated with the development of a fledgling Spatial Data Infrastructure for a commercial archaeological unit in the United Kingdom.

The central component of this approach is the concept of an event driven model. This differs slightly from a more traditional entity based model in that the focus in on the events which produce information objects with the information objects themselves becoming secondary. This in and of itself produces greater semantic clarity as the spatio-temporal nature of existence becomes firmly embedded in the data model. As described in the previous section, our world can more easily be described using inter-related sequences of events.

2.1 Events

Our archaeological information systems therefore need to be built around events. With respect to implementation, we can theorise two broad groups of event. Firstly, those which happened in archaeological or historical time, ie the events we wish to describe, infer and ascribe meaning to. The other group of events is the events which we as archaeologists conduct in order to document, assess, analyse and interpret the archaeological record. In other words, a set of events in the present used to investigate a set of events in the past, ie the first set of events and any associated archaeological remains.

In reality, these events are all the same class of data objects, the key difference between them being their relative position in time. Activities conducted by archaeologists through the process of doing archaeology happen in the present, or in some historical present compared with events being investigated which necessarily happened prior to that point. Furthermore, as time inexorably advances, our own activities will become part of the archaeological and historical record as the records of our forebears have. In other words, it is perfectly possible for the first group of events, ie those being investigated, to be our own activities at some point in the future.

Each event of any type is bounded in space and time and it is this spatio-temporal aspect which is a powerful construct for archaeological information systems. It is this which allows us to use computer based reasoning on our now explicitly recorded events in the past, for example to apply Allen’s Operators to visualise and investigate chronological sequences, and also this which allows us to recognise and record the chains of inference and assertion in the present and their socially constructed nature.

Taking our pot as an example again, we can create a set of records which better reflect the life story of the object.

A pot is manufactured by some person or persons. This involves some physical material and occurs at a location. The material is significant as is where this event takes place which we may or may not know for certain.

The pot is used and undergoes many events through the course of this use period. Some of these use events may leave residues, more physical material, on the pot which may tell us more about particular use events such as where or when they took place and what substances or materials were involved.

Eventually, the pot or parts of it end up in the ground through some kind of depositional event. Through a range of transformative events the pottery eventually finds itself in the archaeological context from where it is eventually found...
...and it is through this finding event that we know of the pot and how it came to be excavated and recorded, through related excavation and recording events.

The pot is then subject to a range of conservation, assessment and analysis events which all add to our knowledge of the pot and the records of which can then be used through subsequent interpretive events to infer information about the originating excavation context.

Looking at our events in the past, we can record each one in more or less detail, depending on availability of data and we can infer that some unknown events must have taken place and at which point in the overall sequence of events. This is very similar to that way in which traditionally archaeologists piece together the archaeological record, the difference here is that the data structure for the archaeological records is more closely allied with this thought process.

Looking at our events in the present, especially from the perspective of managing large fieldwork projects with numerous teams of specialists and interdependencies between strands of investigation, having access to the chains of inference and being able to track change through the system by means of sequences of events is of significant advantage. By allowing multiple assertions as the products of explicit events, it is possible to hang management triggers off the archaeological records rather than having project and records management largely divorced from the process of knowledge creation and compilation. In particular, our events can act as triggers for other events relating to revision and re-examination. So, for example, the addition of a new set of interpretive data through some kind of eg classification event or assignment event (such as period ascription, finds classification and/or dating, etc) can be used to propagate data and change notifications through to other specialists to whom this is important; updates to the stratigraphy can easily inform finds assessment/analysis and vice versa, likewise with environmental data or indeed any other form of data. Whilst we have introduced multivocality into the information system, arguably making the knowledge base more complex, we have at the same time introduced mechanisms through which version control and currency of information can be much better managed, by being able to see and use the chain of events and be explicit about which point in this chain any particular event took place.

2.2 People

The remainder of the basic modelling elements are the places, people and stuff. If we begin with the people, it is true that whilst it is the people in the past who we wish to understand and who’s activities we investigate, it is rare that we can ever really talk about individuals in our archaeological narratives. We know that the pot described earlier was the product of some kind of manufacturing event and we can infer that that particular event involved one or more people, but we know little about these people other than at the time of the manufacture of the pot, they were present. We know these people existed and ultimately we wish to know these people, but in most cases we only get to know about the events in which they participated and the physical remains they left behind.

Of arguably prime importance with respect to the class of data objects representing people are therefore we as archaeologists, for it is us who create the records, write the narratives and ascribe meaning to the remains we find and objects we study. It is us who describe the past and recreate those past people through our investigations. This process is not neutral and is biased by the availability and choices made with respect to data not to mention social context, political agenda(s), etc. So a key part of an archaeological record is not just the bald observations, the metrics, the tables and charts but the metadata about the creation of these records; the who and the why.
Taking our pot example once more, the pot will ultimately be quantified and classified by an expert. In some cases, Roman amphorae for example, there are multiple classificatory schema; one specialist will use of typology whereas another will use another. Chronological sequences based on the known evidence surrounding these typologies is subject to change as new evidence emerges. Currently, the downstream impacts of revisions to typologies are difficult to track and manage but by explicitly storing not just a period ascription for an excavated feature but the salient components of the inference chain, these impacts can be much better managed; the assigned date range for an excavated feature therefore becomes a record that a particular specialist assigned a particular type to our pot, a type derived from a particular typology maintained by a particular specialist or specialists and ultimately this type assignment was used in association with the type of deposition event to inform an assignment of a date or date range to the original excavated feature. Currently, the bulk of this inference chain would only ever exist in the head of the person writing the narrative or at best as a set of notes surrounding the narrative, thus making it much harder to track change through any information system.

Crucially, this same mechanism allows for true multi-vocality. If we are explicit about which archaeologists and specialists created which records, based on which evidence, then we are not forced into accepting a particular viewpoint because the information system cannot support anything else. It is accepted that this can lead to complex user interface issues and problems regarding user experience with any information system but to gloss over the complexity of archaeological information is less than satisfactory; if one strand of evidence if telling one story and another strand telling a different story, then there needs to be a discussion about that with data regarding that interpretive event and any outcome(s) recorded.

2.3 Places and Stuff

Places and stuff are arguably the easiest to define. These concepts are well used and well understood in archaeology. We use survey instruments to record locations and we use named places to record less well defined locations. We use coordinate systems of varying types in both two and three dimensions. Through our activities we engage with physical remains be they deposits, objects or structures.

The key difference being proposed here is that records pertaining to places and stuff become the products of events. The definition of a context boundary (ie a place) is the result of an interpretive event whilst the physical manifestation being observed and recorded is the product of one or more events, potentially a complex sequence of anthropogenic and environmental events. Classifications of objects and structures are not inherent properties of those objects and structures, rather they are the products of interpretive events. So we are using the same processes for observation and recording but the recording system into which these records are created better reflects the way in which reality occurs and how we (implicitly) think of the world around us.

2.4 Realisation

In practice, how this conceptualisation moves through to a logical and ultimately a physical representation using some kind of database platform is dependent on the usual range of factors as with the design of any information system; conceptually though, the fundamental of having two sets of events, one in the past and one in the present, linked by place and archaeological remains, is the key and is platform independent.

It must be stressed that such approaches to data are of particular importance with respect to re-use and re-purposing of data. Semantic clarity is vital for effective search and retrieval, automated processing of data
and the interoperability of systems to name but a few scenarios and, given the nature of archaeological data as described earlier in this paper, efforts must be made to harmonise our shared resources in some way (Binding et al. 2008). This can either be through mediation using ontologies such as the CIDOC CRM (Doerr, 2003) or by using robust, semantically clear, robustly modelled data structures or a combination of both.

To date, digital archives such as the Archaeology Data Service have made archive material available in the form of tabular datasets with supporting metadata, which to some extent achieves clarity in that it is beholden upon the end user to ensure they fully understand a dataset obtained. This understanding comes through rebuilding the separate tables and using the supporting documentation to reconstruct an information system. As we move towards more dynamic modes of dissemination, with archives becoming available as live resources to be shared, consumed and ‘mashed up’, the importance of semantic clarity integral to the data becomes paramount; consuming disparate resources for research purposes using Linked Data approaches where it is not clear what the data means potentially has dramatic impacts on any research so conducted. Given that Spatial Data Infrastructures and more complex archaeological information systems are starting to appear and gain broader adoption, adopting a suitable underpinning approach towards semantic clarity is essential, as will be discussed in the next section.

3. Putting it all Together

This section describes how a UK based commercial archaeological contracting unit is making use of these techniques to capture, store, analyse and ultimately disseminate archaeological data. It has been appreciated that given the scale of some of the fieldwork operations regularly undertaken, digital archaeological information systems are a logical choice for working with data (Cripps, 2012).

3.1 Background and Context

Databases have been in use for a considerable amount of time with GIS becoming widely adopted over the last ten to fifteen years. Experiences gathered through participation in projects such as the Framework Archaeology projects have also fed into a programme of systems development.

The current position is that more integration is required between the various databases/GIS in use. Tighter linkages to core business processes (administrative, financial and marketing/customer relationship management) would also be advantageous. In addition to this, drives within archaeology towards making more data available in addition to narratives produced as books and reports are having an impact on requests and requirements for data supply to clients, curatorial teams, archival repositories and the general public.

To this end, Wessex Archaeology has started to develop a Spatial Data Infrastructure (SDI) based approach capable of supporting these functions. It will be event driven taking into account the results of a recent process mapping project which looked at team functions and inter-relationships. Using a combination of proprietary and open source components, the SDI comprises a number of linked modules, each designed to fulfil a particular purpose. Of particular relevance here are the output functions which will facilitate the publication of Linked Data compliant with the CICOC CRM by means of the STELLAR toolkit using RDF as the exchange format. This will allow the data to be used alongside other datasets published in this way, a growing list, and offers significant advantages over simpler modes of publication such as unmediated Linked Data or a series of tables, as described above. Input functions are also vital and some discussion will be afforded the currently available mechanisms for receiving data.
So, internally within the organisation, data will be gathered into robustly modelled databases from where it can be used, reused and enhanced. There are many and varied requirements here including but not limited to use across multiple sites by multiple teams, use in the field, use by external specialists. Each of these will be accommodated through a combination of technology and supporting protocols, help and training. A modular approach allows each component to be built and maintained separately and replaced as needed with any particular architecture developed to suit the specific needs of a component. The overarching standards compliant framework facilitates outputs to digital archives repositories and other third parties, subject to any applicable licensing constraints. Throughout, the aim is to allow data to flow easily, be semantically clear and be enhanced and enriched on its journey through our archaeological process before ultimately flowing into a recognised digital archive repository ready for further use and analysis.

A major obstacle to the seamless flow at present is the nature of incoming data. Data supplied from local Historic Environment Records (HERs) or the English Heritage archives and repositories can be problematic as even where there are standards in place, local implementations of these can vary. Furthermore, data is typically supplied as static snapshots which can be more or less complete. Finally, supply from multiple sources may not correlate. All of this is a direct product of the way in which heritage data in the UK is managed through a disjointed and disparate collection of groups and organisations.

Presently a lot of time and effort is devoted to exchange and recompilation of databases. In the same way as this was recognised as a problem within Wessex Archaeology, it is now time to recognise that the way in which we manage heritage data in the UK more and further afield is problematic. The UK situation comprising multiple tiers of records held across organisations combined with receipt and supply to and from multiple record systems simply duplicates resources, wastes time and money, and creates problems with data quality and consistency.

3.2 Moving Forward

Steps are now being taken to remedy this with numerous archives becoming available as Linked Data. This is a great way forward as it enables improved access to resources and (in many cases) the process of creating the linked data has ironed out any semantic differences through the use of CRM-EH as a mediator.

The Heritage Gateway system now aggregates content from the tier of archaeological records known as Historic Environment Records or Sites and Monuments records plus participating national archives although no effort is made here to harmonise the data in any way, it is up to the end user to make sense of any data supplied. Commercial use of and access to such resources is also prohibited.

We currently do not have in place robust protocols or mechanisms for exchange of data to and from archaeological contractors which impacts on any developments being made here by said contractors. Furthermore, exchange of data between the various local and national inventories is similarly problematic, hence the lack of robust cross-referencing and other inconsistencies between their holdings. Indeed, much exchange of data is still undertaken using paper based mechanisms or digital mechanisms which simply replicate existing paper based mechanisms (for example database reports as pdf documents).

Overall, there needs to be a move towards a more integrated system supporting a single national archive. I would argue that this should be a national but distributed information system to realise cost savings based on scale; simply reducing the number of servers from
over a hundred to one will bring significant savings. Costs relating to development and support could also be reduced, following the model currently used by around half of UK HERs using the HBSMR system from exeGesIS, whilst implementation of more complex components such as MidasXML routines, the English Heritage XML based data exchange protocol, would be more possible as issues relating to the capacity of host organisations would be eliminated. In addition to cost savings, a key factor in these austere times, such a system would bring significant advantages to users through better integration of currently divorced datasets and improved exchange mechanisms, better supporting the commercial archaeology sector and research projects; extended periods allocated for collation and integration of data at project initiation could be dramatically reduced. Importantly, any such system, whilst being centrally developed, hosted and supported as a unified entity, must be managed locally, reinforcing the position of those with the best local knowledge. Obstacles here are now largely political rather than technological so it is very much hoped that the systems emerging will pave the way for broader integrative Spatial Data Infrastructure approaches across the discipline and outside to other related disciplines.

3.3 Conclusions

The development of an SDI within a UK contracting unit is aimed at improving the way in which archaeological (spatial) data can be captured, managed, stored, analysed, archived and published. By using event based modelling as the basis for this development, we will be able to better handle the necessary subjectivity and multivocality present in archaeological research, by explicitly recording the personal involvement of archaeologists. This will also assist with version control and currency associated with the necessarily iterative processes of post-excavation assessment and analysis, taking into account uncertainty and error present in archaeological data. The intention is that this new approach will further facilitate the processes of archaeological assessment and analysis by providing a platform upon which digital interfaces can be built, notably to interact with the digital site archive through temporal and spatial visualisations such as dynamic stratigraphic matrices and schematic drawings. It will also facilitate spatio-temporal investigation by exposing the data to automated processing and potentially external systems.

Extrapolating this approach out to the broader heritage sector promises to bring further benefits. Academic research can draw on the same core elements regarding subjectivity, multivocality, temporality and uncertainty. Curatorial and management programmes and responses to planning consultations as part of development control can be better provisioned through more integrated resources whilst research agenda and research frameworks can truly become data driven and conversely feed more directly into these other aforementioned areas.

So, in conclusion, the combination of ontologically enabled, event based models for archaeological information systems combined with ontologically enabled, spatial data infrastructures is now technological feasible,
offers significant benefits and takes us one step closer towards a spatially aware semantic web for heritage.

References


ArcheoInf, the CIDOC-CRM and STELLAR: Workflow, Bottlenecks, and Where do we Go from Here?

Geoff Carver
Universität Göttingen, Germany

Abstract:
The following reports on work undertaken to implement the CIDOC-CRM for the ArchaeoInf project. ArchaeoInf (www.archeoinf.de) is intended to act as a data repository, a database which combines and integrates primary archaeological data collected and processed during surveys or excavations and originally stored in individual project databases. Although creating ArchaeoInf presented a number of difficulties (many related to terminology), the following focuses on an aspect of the creation of metadata mapped to the CIDOC-CRM for export in RDF format using the STELLAR “data mapping and extraction utility.” The “standard” STELLAR templates enable exports into the English Heritage version of the CIDOC-CRM, but since ArchaeoInf is not concerned with databases or projects using English Heritage documentation standards, modified versions of the “user defined” templates were used instead. This paper describes the process of mapping existing databases into the CIDOC-CRM and then creating STELLAR templates for exporting the results of SQL queries into an XML/RDF format.

Keywords:
CIDOC-CRM, STELLAR, Database Mapping, “How to?”

1. ArcheoInf, the CIDOC-CRM and STELLAR

This paper describes the process used to prepare data for export from the ArcheoInf Database (www.archeoinf.de) using the STELLAR “data mapping and extraction utility” (Tudhope et al. 2008, 89) (http://hypermedia.research.glam.ac.uk/resources/STELLAR-applications). It is intended to provide a straightforward, “how to” outline to supplement any number of technical papers already available. The author developed the basic process while helping to map the Arachne Database for the Universität Köln (http://en.wikipedia.org/wiki/Arachne_%28Archaeological_Database%29). Arachne is described as being “a central object-database for a large federal institution which possesses about two millions of images inside their photographic archives and produces even more data each year in the course of its research activities” (Wikipedia 2012).

At the time of the author’s involvement during the winter of 2008-2009, Arachne consisted of ca. 125 tables originally prepared by the various offices of the German Archaeological Institute (Deutsche Archäologische Institut/DAI) in Rome, Athens, Istanbul, Cairo and Madrid. The various tables document artifacts, photos from photo archives, bibliographic references, etc.

To put these numbers in perspective, Arachne might be compared to the 60 tables, ca. 2500 fields, ca. 250 index lists and ca. 110 layouts of iDAI.field (DAI 2012), the DAI’s database for recording sites and field surveys. ArcheoInf itself has 46 tables and a total of 225 fields.

Following a feasibility study by Robert Kummer (2007), it was decided to map the contents of Arachne in CIDOC-CRM (CRM) format. It was later decided to also map the contents into Dublin Core (DC).
Since the large number of tables and fields did not allow an easy overview and Arachne itself had not been documented, the first step was to prepare an inventory of the tables and their fields. An inventory sheet (Table 1) was prepared for each table listing each column heading (exported as a list from the database) plus scope notes, data format (integer, text string, Boolean, etc.), examples, comment, and the CRM and DC mappings. Each inventory sheet also had a link to a PNG version of a flow-chart which illustrates the mapping.

When examining the contents for the purpose of writing the scope notes and comments, no attempt was made to correct errors within the actual content of the database. Examples were chosen which not only showed what kind of information was intended to be recorded within a given column, but also included examples of entries which clearly did not belong. In addition, since many columns were empty, their function had to be assumed, often on the basis of very unclear column headings.

The overall situation was similar to that encountered during the course of the STAR project:

*In most cases we did not have specific metadata nor descriptive scope notes on what the field content was intended to represent. It was therefore important to review the actual data within the field, as well as the field label, to judge the intended ‘meaning’ of each field. We recommend assessing the data content, even when a field label seems familiar (Tudhope et al. 2011, 2.2).*

In the case of Arachne, the lack of documentation reflects the context (multiple locations) within which data was entered (Carver and Lang 2013).

These inventory sheets were linked to a master document which could in turn serve as the documentation for the Arachne database as a whole. These inventory sheets and the master document were prepared using Open Office Writer. One clear advantage of Open Office over, for example, Microsoft Word, was the ease of producing PDFs without the need for additional software or installing a “printer.” Thus an updated version of the master document was produced almost daily with the intention that the final version would eventually be posted online in PDF and XML format.
These inventory sheets were intended to provide the basis for some form of template which would in turn be used to export data as XML tagged with the relevant CRM and DC mappings suitable for downloading in response to SQL queries, in addition to displaying an SVG form of the flow-chart online. The process for creating these templates and export documents had not yet been determined when the author quit the project.

For the purposes of the CRM, each individual table was considered a document as part of a larger document (the Arachne database itself). Each table thus documented a CRM “Entity”: a place, artifact, site, image, etc.

2. CRM Mapping

Mapping was done by consulting PDF versions of the Definition of the CIDOC Conceptual Reference Model (Crofts et al. 2009; Tudhope et al. 2011, 1.2) and the online Web Ontology Language (OWL) version maintained by the Friedrich-Alexander-University of Erlangen-Nuremberg (http://erlangen-crm.org/) viewed in Protégé (http://protege.stanford.edu/). The advantage of using the PDF version reflects the ease of searching electronically rather than leafing through a long printed document, while the online version in Protégé enables exploration of the hierarchical relations between Entities graphically.

The CIDOC-CRM is a powerful formal (and occasionally infuriatingly abstract; Tudhope et al. 2008, 88) metalanguage. Its logic is not immediately apparent, but may be illustrated through an example of one of the many conceptual problems the author has encountered using legacy data from a cemetery in Italy (Heitz 2009). Since the human remains in this case had either decayed (after burial) or could no longer be located (after excavation), there was really no “person” involved (i.e. they were only documented indirectly), and the documentation should have centred on graves as “Man-Made Features” (i.e. holes in the ground). As a grave itself does not have a gender (in some cases this could be inferred on the basis of grave goods or the size and/or depth of the grave [assumed to reflect status]), nor an age class (juvenile, adolescent, adult) – despite not actually having been documented in the database itself – the “person” had to be modelled in the CRM.

At the same time, the grave also had to be modelled. In some cases, for example, although no body had been found, there was evidence of a “crouched” burial. And just as a grave cannot be “adolescent,” the person who had been buried there was not “crouched.”

Partly in order to develop methods and the experience necessary for dealing with such problems, it was decided to begin with simple cases and gradually work towards more complex mappings. A number of tables which simply listed “internal” links to other tables were quickly mapped and the knowledge gained used to provide a work flow and overview of the way forward. In this way, duplication sped the mapping: sections which had already been mapped for one table could be reused by simply copying and pasting relevant strings of text. Since each table, for example, was considered to be a “document” documenting a given “CRM Entity” (i.e. photo, book, architectural fragment, statue, etc.), and each “CRM Entity” had, in turn, a “Preferred identifier” (index or primary key) which identified it within the database, these “roots” could then be copied and pasted into each inventory sheet.

After this, relatively easy (and often repetitive) elements were identified and combined like building blocks. The process of constructing Arachne by combining existing tables from a number of sources had led to a good deal of duplication. Information on places was repeated on any number of different tables, for example, as the location of sites, places where stray finds had been discovered, books published, places depicted in photos, etc. In
many cases the tables recorded who had last edited a given record and the date when the data was entered, and this data was linked to the table or “document” itself rather than to the entity recorded. After such common links had been identified and modelled once, they were simply copied and recombined.

Table 2. The script used to generate the Graphviz diagram in Figure 2.

| digraph g { | Indicates the beginning of a “directed graph” (opens brackets) |
| graph [ rankdir = “LR” ]; | Declares that the graph’s layout (“rank direction”) will be Left to Right (i.e. not vertical) |
| node [ fontsize = “16” shape = “ellipse” ]; | Gives information about the format the “nodes” will take: ellipses with 16 point text |
| “E84” [ label = “<f0> E84.Information_Carrier | The sculpture ‘Monument to Balzac’” shape = “record” ]; | Defines an entity called “E84” with a label “E84.Information_Carrier” (CIDOC-CRM entity) in a box separated from “The sculpture ‘Monument to Balzac’” (in the horizontal layout, the vertical line placed the text under the CIDOC-CRM identifier); the shape is defined as “record” (i.e. a box) |
| “E21_Balzac” [ label = “<f0> E21.Person | Honore de Balzac” shape = “record” ]; | Defines an entity called “E21_Balzac”; the <f0> defines an address in this box (in this case the portion containing “E21. Person”; other lines could be similarly numbered to allow more complex links) |
| node [shape=ellipse]; {node [label=“P62. depicts”] P62;} edge [ ]; | Defines a relation (node) called “P62” and labelled “P62. depicts” |
| “E84”:f0 -> “P62” [ id = 0 ]; | Links “E84” with “P62”; note the direction of the arrow |
| “P62” -> “E21_Balzac “:f0 [ id = 0 ]; | Links “P62” with the relation “E21_Balzac”; |
| label = “Monument to Balzac”; fontsize=20; | Provides a label to the directed graph |
| } | Indicates the end of the “directed graph” (closes brackets) |

3. Graphviz

GraphViz (http://www.graphviz.org/) was used to render the CRM mappings graphically. The original aim was to provide a means for controlling the mappings as they developed (correcting errors, etc.). Within the inventory sheets the column headings were listed according to their “sequence” within
the various tables in order to ease interaction with the database, but with some of the more complex tables this was far from ideal, as the columns were often listed alphabetically, without consideration for logical or semantic integrity (i.e. “collocation” or grouping closely related columns together; Beall 2006, 62). Although the column headings could have been reordered within the documentation to better reflect the semantic structure, it was quickly recognised that the resulting diagrams also helped document the various tables.

When choosing a graphics program, the main concern was that the layout should be generated automatically. Anyone using such flow-chart programs as Visio (Visio 1996) generally has to arrange various elements. As the goal was to find a way to quickly check the logic and internal consistency of the various mappings, so long as the final drawing was comprehensible, the layout itself was relatively unimportant.

Since GraphViz is also used in the Protégé ontology editor, consistency in layout and coding of output files (OWL, RDF) was seen to be an advantage, as was its ability to export diagrams in SVG format for use online. A number of plug-ins are available for GraphViz which can, among other things, extract and automatically graph a database structure.

The STELLAR programmers recommend the use of a similar program called Gruff (http://www.franz.com/agraph/gruff/) as a means for querying the resulting RDF. Use of one or the other may simply be a matter of personal preference.

GraphViz has the benefit of being freeware, and is relatively simple to use. Various objects and their relations are defined as text, meaning that the process can be automated to some extent by exporting a list of the column names and using a universal search and replace function within a text editor2.

The column headings were exported and copied into the text-editor. In this way, complete lists could be turned into drawing elements quickly, without error, and those entities which had not been linked up in some relation were immediately visible.

Each “Entity” was mapped in a box (shape = “record”). Those Entities which corresponded to a given column were drawn to show the column name in a lower box. The Entities were linked through “Properties” mapped as ellipses (shape = “ellipse”).

Table 2 is a sample of the script used to draw a simple Graphviz diagram (Fig. 2; Doerr and Kritsotaki 2006, 5 and fig. 3)).

The mappings were drawn horizontally, leading from the name of a given table (at the far left) to the name of a given column at the end of each branch (Fig. 1). Although the diagrams could also have been generated vertically (top to bottom), the results were less consistent (and more difficult to read). Among other things, when a number of fields (the various spatial coordinates of Cartesian space, for example, or length, width, thickness or diameter and weight of a given artifact) all mapped onto the same path (thereby requiring a list of column headings), the lines joining individual elements from any given list become especially complicated in a vertical format.

2 The author used TextPad (www.textpad.com) largely because it allows the simultaneous editing (search and replace) of numerous documents.
As has been noted, each table was considered to be a “document” which documents a given “CRM Entity” (i.e. photo, book, architectural fragment, statue, etc.). Each such document was given a CRM “root”: the table (E31.Document) documents (P70) photos (E38.Image). These “roots” were marked in gray to distinguish them from the contents of the table themselves, thus allowing the activity of entering data to be linked to the table/document and not to the entity being documented.

This “P70” solution actually caused some debate. Some of the computer programmers insisted that an Entity is documented in (P70i) a given table (E31), but this caused problems with those tables which only combined “internal” links to other tables, and thus did not document anything per se. It also caused conceptual problems when trying to link the individual tables back to Arachne, since Arachne – as a Document – is composed (P148. has_component) of parts (i.e. the individual tables, each of which is in turn a Document), each of which documents (P70) some type of CRM-Entity (E1). Otherwise, if an Entity is documented by (P70i) a given table, then Arachne documents a type of Entity which is in turn documented in a table (Document). As a database, Arachne does not document objects; Arachne merely links a series of tables, each of which documents a type of object in the real world. These tables can exist separately from Arachne, and may be added or deleted at will.

Some of the entities which were recorded in two or more tables were not distinguished in the CIDOC-CRM. The database records artifacts. Some of these are sarcophagi, recorded in the “sarcophagus” table, and some were statues, recorded in their own table. It is simply easier to divide the mapping according to existing tables and thus maintain consistency with the database itself, especially considering that it was continually being modified.

It was also intended to eventually make it possible to link each table diagram in a diagram of the entire Arachne database. This would have required some minor editing (fixed when it came time to map ArcheoInf): instead of using just the column name as an identifier (“ID_number”), the “entity” nodes were identified

<table>
<thead>
<tr>
<th>Alter</th>
</tr>
</thead>
<tbody>
<tr>
<td>E31.Document</td>
</tr>
<tr>
<td>P70_Documents</td>
</tr>
<tr>
<td>E21.Person</td>
</tr>
<tr>
<td>P2_has type</td>
</tr>
<tr>
<td>E55.Type</td>
</tr>
<tr>
<td>txt_alter</td>
</tr>
<tr>
<td>P48.has_preferred_identifier</td>
</tr>
<tr>
<td>E42.Identifier</td>
</tr>
<tr>
<td>ID_Alter</td>
</tr>
<tr>
<td>E42.Identifier</td>
</tr>
<tr>
<td>P48.has_preferred_identifier</td>
</tr>
<tr>
<td>E55.Type</td>
</tr>
<tr>
<td>P2_has type</td>
</tr>
<tr>
<td>E21.Person</td>
</tr>
<tr>
<td>P70_Documents</td>
</tr>
<tr>
<td>E31.Document</td>
</tr>
</tbody>
</table>

Table 3. Example of a completed CRM map.
using a binomial combination of “table-name” and “column-name” ("person-ID_number") to insure that all the elements from each table were uniquely identified. In this case, the larger E31.Document “Arachne” would have all the separate tables as parts.

4. **Arachne Export**

Once each inventory sheet had been prepared, a copy was made where all of the extraneous information (scope notes, examples, etc.) was stripped away, leaving a simple version listing just the CRM mapping and column names, ready for conversion into XML (including the use of tabs to mimic the hierarchical structure (Table 3)). The column order was also rationalized (“collocation”), reflecting the CRM rather than the alphabetic or apparently random sequence found in the database itself. Copies were also made of these secondary files, where the CRM maps were reduced to DC (Table 4).

Scope notes, SKOS, etc., could have also been reintegrated into the list at this point, if desired.

5. **Lessons Learned from Arachne**

One of the most important (and obvious) lessons learned from working with Arachne (and reinforced by subsequent experience working on the ArcheoInf project) is that databases (and software) need to be documented (Backhouse 2006: 54-55), as this saves a great deal of guesswork later when trying to determine what a given column was meant to represent. Although this need for documentation may be obvious, corners are cut too often. An especially difficult case was that of a table where the column headings correspond to a large number of the 650+ numbered (not named!) fields of the MARC 21 bibliographic standard (http://www.loc.gov/marc/bibliographic/).

It also helps if the database itself has been well-designed. The spatial information stored in Arachne provides a case in point. Ideally, one table would contain the name of a place, its local jurisdiction (county, province, etc.), which country it lies in, plus spatial coordinates and perhaps a cross-reference to a map reference and/or an online gazetteer (such as the Getty Thesaurus of Geographic Names® Online (http://www.getty.edu/research/tools/vocabularies/tgn/)). As has been noted, spatial information was entered repeatedly on numerous different tables. Given the circumstances under which Arachne was put together in different locations, this probably could not have been avoided, but it was not efficient. More important, perhaps, is control over data entry, to insure that only one type of information is entered into any given field: only Spatial Coordinates (E47) and not town Place Names (E48), for example.

Experience with mapping Arachne in CIDOC-CRM shows the value of beginning with something simple and working up to more complexity (Woolley 1961: 54). The CIDOC-CRM is complex enough without starting at the most complex table possible. Starting with something relatively easy allows someone to find out how the CRM works, become familiar with the various aids (the handbook, ontology editors like Protégé, etc.), and to develop a work flow process that suits the given needs and available resources. It is worth noting that, for example, whereas the interactive interface in the Erlangen OWL makes it easy to follow entities (i.e. E59.Primitive Value, E60. Number, E61.Time Primitive and E62.String) through their hierarchical relationships, some of the elements listed in the handbook are not represented online. And, although having a “hard” copy of the handbook is always a good

<table>
<thead>
<tr>
<th>Alter</th>
<th>CIDOC-CRM</th>
<th>DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>txt_alter</td>
<td>E55.Type</td>
<td>Type</td>
</tr>
<tr>
<td>ID.Alter</td>
<td>E42.Identifier</td>
<td>Identifier</td>
</tr>
</tbody>
</table>

Table 4. CIDOC-CRM mappings reduced to Dublin Core.
The “search” function in the PDF format is probably more helpful in the long run (among other things, the handbook is too long to search through manually).

Potential users really have to learn how to use or think in terms of the CIDOC-CRM:

many of the entities in the ontology are fairly abstract; understanding the conceptual complexity of the CIDOC CRM poses a challenge to some non-specialists. It can also be possible for different people to make alternative valid mappings to the ontology for the same situation, raising difficulties for semantic interoperability (Tudhope et al. 2011, 4.2).

Eventually, with practice and familiarity, it is possible to model something as initially intimidating as a list of the items carried by a given deity depicted on a certain portion of a coffin in the CIDOC-CRM (“E22.Man-Made_Object” – “P58.has_section_definition” – “E46.Section_Definition” – “P62.depicts” – “E1.CRM_Entity” “P3.has_note” – “E62.String”), but to figure out how to do this efficiently takes time and practice, as was underlined when attempts were made to map several very complex tables “out of sequence.”

On a practical note: when working simultaneously with Open Office Writer, a text editor, Gruff or GraphViz, the online Web Ontology Language (OWL) version of the CIDOC-CRM viewed in Protégé, and a PDF version of the Definition of the CIDOC Conceptual Reference Model, it is useful to work with two monitors.

Since some specialised software will not run on all operating systems, hardware and operating systems also play a role. While working on Arachne the author had one machine running Mac and another running a Windows emulator.

### 6. ArcheoInf

The purpose of ArcheoInf differs from that of Arachne. As described elsewhere (Carver, Lang, and Türk 2012, Carver and Lang 2013), ArcheoInf is intended to function as a long-term document repository for combining and storing primary archaeological data recorded in individual project databases constructed for use on surveys and excavations undertaken by German university departments of Classical Archaeology. As such it might be compared to the ADS in the UK (http://archaeologydataservice.ac.uk/), or DANS in the Netherlands (http://www.dans.knaw.nl/en) or tDAR in the USA (http://www.tdar.org/). The major difference being that, whereas the ADS stores documents, ArcheoInf aggregates data from any number of databases in a new database, thereby allowing inter-site queries. A closer analogy might be the STAR project (May et al. 2011).

Funded by the Deutsche Forschungsgemeinschaft (DFG), ArcheoInf was developed as a collaborative project involving the following institutions:

- Archäologisches Institut und Sammlung der Gipsabgüsse der Georg-August-Universität Göttingen;
- Institut für Archäologische Wissenschaften der Ruhr-Universität Bochum;
- Fachbereich Vermessung und Geoinformatik der Hochschule Bochum;
- Universitätsbibliothek Bochum;
- Lehrstuhl für Softwaretechnologie in der Fakultät für Informatik der Technischen Universität Dortmund;
- Universitätsbibliothek Dortmund.

Building on previous experience, the CRM and DC mapping followed much the same process as that used for Arachne. Where relevant, the columns were also mapped in
SKOS (Lang 2013) and Europeana (Isaac and Clayphan 2011) format.

Because the ArcheoInf database was planned from the beginning and not assembled from existing tables and databases, it is simpler and more coherent than Arachne. On a technical level, however, some of the elements unique to ArcheoInf (the thesaurus and the relations between various concepts) present special problems. The CRM allows for hierarchical relationships between “types,” for example (a sword is a “type of” weapon), but unlike SKOS, makes no allowance for related terms (i.e. near-synonyms or the foreign translations of general terms in the thesaurus which is one of ArcheoInf strengths (cf. Baines and Brophy 2006, 239-240)).

7. Using STELLAR to Export ArcheoInf

It was decided to use STELLAR to export the data tagged with CRM mappings (May et al. 2011). One of the advantages of this software is that it, too, is free. STELLAR includes templates built for use with the English Heritage version of the CIDOC-CRM (CIDOC-CRM-EH) (http://hypermedia.research.glam.ac.uk/resources/crm/), with columns relating to particular fields within English Heritage documentation system, and alternate “user-defined” templates. Since ArcheoInf does not use the English Heritage documentation system, a text editor was used to modify the “user-defined” templates to reflect the contents of the existing tables.

The STELLAR example files were downloaded and their contents examined using a text editor in order to determine the function of each one while working through the examples in the tutorials (http://reswin1.isd.glam.ac.uk/stellar/tutorials/tutorial1.html, http://reswin1.isd.glam.ac.uk/stellar/tutorials/tutorial2.html). This method made up for the fact that the tutorials were generally instructive without being explanatory (focusing more on what to do without explaining why any given step should be performed, a common failing in computer documentation).

One minor criticism: when consulted, some of the weblinks on the STELLAR website were broken. These have since been fixed, but the lack or absence of samples can be frustrating to anyone working to deadline or trying to choose between one program and another, or decide whether to trust software to the degree of investing time and effort getting it to run (Carver and Lang 2013). There are also minor formatting issues with the output (a lot of unnecessary blank lines), and some problems opening the results in Protégé (possibly some formatting incompatible with RDF or OWL).

Another minor problem reflects the fact that, like many other projects, STELLAR syntax is dependent upon imported code. Those seeking “more details of syntax and more advanced usage examples” – such as information on the use and constraints on some of the “if” statements – are directed towards the site of StringTemplate (described as “a java template engine”, http://www.stringtemplate.org/), which, like many such websites, is not particularly “user friendly.”

Of more concern, though, was the fact that, as a terminal-application, the STELLAR user interface was not the easiest to use; typing mistakes are inevitable. A new, user-friendly GUI has since been added for use in Windows. The process was worked through by experiment first with carried out simple CSV tables (Klammt 2011), then a slightly more complex Access database (Heitz 2009), before starting to work with ArcheoInf itself using SQL queries. As has been noted,

In practice, getting to use someone else’s data implies mapping the fields and codes used to one’s own system. This is a time-consuming exercise, that has to be repeated each time a different data set is imported. And even then, data that cannot be adequately translated may be lost in the process (Verhagen et al. 2011, 153).
The data first had to be formatted properly. The CSVs exported from MS-Excel caused problems (probably due to the inclusion of “illegal” – and invisible – formatting characters), so were re-exported from Open Office Calc. The German “sz” (ß) and letters with umlauts (ä, ö, ü) were also converted to the standard character set (ss, ae, oe, and ue).

The next step will be to put the query and export process function online, either using STELLAR or some alternative. Further development has been limited in part by the need to acquire data (Carver and Lang 2013).

8. Conclusions

Overall, the process works. It is possible to map complex archaeological databases and export tagged data with the CRM mappings in XML using STELLAR templates. What is less clear at this point is how this tagged data can then be imported into other databases (i.e. how to “match” mappings) and perhaps whether or not there is even a demand for the ability to do so. At present it may be possible that extensive effort is being invested into export functions nobody wants. The question then becomes one of whether or not that investment represents a waste of time and effort, or whether – once widely available – tagged data will create its own market (Tudhope et al. 2011, 4.6; Eckkrammer et al. 2011, 84).

This paper was intended to help make the mapping process just a little – it is hoped – easier for beginners.

References


Abstract:
This paper examines a range of issues encountered when formulating archaeological research questions and attempting to gather or create Linked Data to address those questions. The paper has two main parts. The first part presents a case study on the kinds of processes by which research questions are generated and structured, based around location and events. It begins with the recent St Pauls churchyard occupations in London and develops related research questions about Boudican period destruction. It describes practical difficulties encountered when trying to gather data to answer those specific questions and raises more general issues of whether existing data is suitable, and available, to answer such research questions using semantic technologies. The second part of the paper examines broader issues raised by the research question driven data investigation and considers what the implications might be for trying to tackle new sorts of research questions using semantic web based technologies.

Keywords:
Semantic Technologies, Research Questions, Linked Data, Semantic Web, CIDOC CRM

1. Introduction

This paper derives from a session entitled “Joined-Up Data: What are the new research questions?” The intention of the session was to consider whether the emerging semantic web technologies actually enable new research questions to be asked, and if so, what could such questions look like and how might they be addressed? The aim of the session and this paper was always intended to be more about understanding the processes by which new types of questions might emerge, along with considering what the new technologies might offer to address such questions. The primary aim of this paper is therefore to consider whether new semantic technologies might help address new research questions, while using a case study example to illustrate issues encountered when developing such research questions.

The paper is therefore presented in two main sections. The first is a case study that attempts to show from a practical example what issues emerge in attempting to consider a specific research topic using semantic technologies to either create or re-purpose linked data. As a consequence, preliminary consideration in the paper has been given to what sort of approaches are taken to initiating and formulating research questions that could be addressed using semantic technologies. The second half of the paper has the main thrust of the discussion and considers some of the key issues that need to be addressed on a much wider scale than just for answering the case study research questions and develops further ideas for how semantic technologies might be employed in the historic environment sector more broadly.

2. Case Study Giving Background to how Appropriate Research Questions might be Generated and Structured in order to be Addressed by Semantic Technologies

The occupation of the churchyard at St Pauls Cathedral began on 15th Oct 2011. This followed a relatively unsuccessful attempt by protestors to target the City of London financial centres for occupation during that day. In the end the protestors set up a temporary
encampment on the areas of pavement around the north-west side and at the west end of St Pauls Cathedral.

At the time of first submitting the abstract for this paper to CAA2012 the occupation was still ongoing, but on 28th February 2012 the occupation was forcefully evicted from the St Pauls area. Another protest site, outside the City of London jurisdiction, at Finsbury Square still continues, at the time of writing this paper, as an encampment of tents and temporary structures on an area of public ground in the centre of an urban square surrounded by office buildings. One of the most evocative images (at least for archaeologists) of the occupations was taken when a vacant office block called “Roman House” on the site of the former Roman Londinium city wall, was also briefly occupied and images of the occupied Roman House appeared in the national press and on the BBC website (http://www.bbc.co.uk/news/uk-england-london-16665217 dated 21/01/2012).

The first kinds of research questions that these and various related events posed to the author was,

- What archaeological evidence might remain from such protest events after 10, 100, 1000, or 2000 years?

- How might that evidence compare with evidence for other protest events, either at, or in the vicinity of the same site?

- What degree of archaeological evidence survives for known sites of protest in the available archaeological records?

This suggested the potential for investigating what archaeological data existed from such contexts recorded on investigation events carried out around the City of London. In addition to the potential for cross-search of data from sites in London there was also the more ambitious possibility of trying to gather data from the two other most famous settlements of the Roman period that were attacked by the Boudican revolt, namely Camolodunum, on the site of modern Colchester, and Verulamium now adjacent to the current town St Albans.

![Figure 1. Modern towns of London, Colchester and St Albans that featured in the Boudican Revolt, along with other prominent Roman settlements.](image_url)
2.1 Case Study: Examples of Research Questions Derived from Relationships between Occupy London site and Boudican Revolt

A preliminary assessment of the potential of the archaeology in the vicinity of the modern St Pauls Occupy London site suggested a number of possible research questions that might be pursued. In particular, given the context of a session for CAA2012 on the potential for semantic technologies and Linked Data to address new research questions, the following questions seemed relevant:

- Where are the excavated sites with Boudican activity in the modern St Pauls area?
- What evidence survives for occupation, or destruction activity, of Londinium, Camulodunum or Verulamium by anyone in the period spanning the Boudican revolt and aftermath (circa 60-62AD)?
- Can we compare archaeological data from projects carried out in the 3 different settlements – Londinium, Verulamium, Camulodunum?
- Can we search on Who is in the record? What issues arise simply when looking for records associated with Boudica, Boudicca or Boudican?
- What are the material remains of the Boudican revolt, compared to evidence for a period of destruction?
- What new research questions might be addressed using semantic technologies?

2.2 Case Study: Methodology

The initial approach was to try and gather available data from excavated sites with Boudican material. The ideal intention was to collect digital data in formats that were most readily amenable to conversion to RDF/XML Linked Data formats. This raised a number of avenues of enquiry, but these could be summarized as:

- Investigate National records of excavations.
- Investigate Regional records of excavations.
- Contact organisations that had carried out excavations or held records derived from excavations.
- Literature review for any published data.

It was fairly clear within the parameters of the Boudican period revolt and modern archaeological investigation in England, that the largest number of possible sites with available data was likely to come from modern London. With the constraint of limited time for actually hunting down data (this researcher was employed full-time on other work), it was pragmatically decided to focus on London first and then to seek related data from Verulamium (St Albans) and Camulodunum (Colchester) in the time available.

The principle sources for data are given in the following section along with a summary of the kinds of data encountered. The intention, and expectation, of this research was not to get exhaustive cover of all the potential sites, but rather to gain an indication of what data was, or might be, readily available and to consider what steps were needed to enable conversion of data to RDF and Linked Data formats and what issues this raised.

2.3 Case Study: Current Sources of Information Investigated: London, Verulamium, Colchester

For data from London the main sources were the following:

- Heritage Gateway (http://www.heritagegateway.org.uk/gateway/),
- Greater London Historic Environment
Record (GLHER) – Monument & Event data,

• Excavated sites by various organisations,
• LAARC – London Arch Archive & Research Centre,
• Publications – e.g. No1 Poultry.

The Heritage Gateway returned 6327 results from the Greater London Historic Environment Record (GLHER) for a search on the period 60-62 AD. This appeared promising so was followed up with a specific email to the GLHER requesting more information on what data was available. Following some helpful email dialogues it was decided to focus the research on excavation sites in the St Pauls vicinity, so a buffer search was carried out based on a 200m buffer around the centre point of St Pauls dome. This incorporated sites on modern Watling Street and modern Newgate street as well as sites that were in close proximity to, effectively underlying, the Occupy London protest.

The buffer search by GLHER identified 70 archaeological event records within 200m radius of St Pauls. The output of the search for event sites was a report in PDF along with the 1:2500 plan showing sites against the OS mapping of the modern street layout (Fig. 2).

The GLHER staff were most helpful in carrying out the search for me but pointed out that the records they hold are indexed according to period, so a specific search on the years 60-62AD was not really ‘meaningful’. The search on a narrow date range defaulted to searching for any sites that fell in the broader periods covered by the English national period thesauri, in this case the Roman period with a time span of 43-410 AD. In addition, possibly not all the most recent events had been reported to the GLHER and some sites were still awaiting entry on the HER.

Examination of the PDF report showed that 13 of the sites were described as excavations, but only in 2 cases was there any reference in the summary descriptions of possible 1st century deposits, and only one of these, Lion Plaza, Threadneedle Street (Event ID ELO760) had mention of actual evidence from 60-70 AD.

Even if more positive ‘hits’ had been achieved, the GLHER at best only holds a quite high level overview of what was discovered by an event, rather than records of contexts, finds or other detailed data. At this point it seemed sensible to approach the excavation teams involved and I therefore contacted the Museum of London Archaeology (MoLA) team along with the London Archaeological Archive and Research Centre (LAARC) to see if I could track down more specific data, either from the Lion Plaza site or perhaps other sites with Boudican material that had not yet appeared on the GLHER.

Email contact with MoLA resulted in confirmation that re-usable data would most likely come from the LAARC, but also noted that the recent publication of No1 Poultry site contained some Boudican data from the excavations. Further email exchanges with the LAARC did track down some possible sites with potential evidence from the Boudican period.

The LAARC has an online catalogue with the ability to search for specific project details
and finds. With the help of the archive staff, between us we were able to return 39 records resulting from a search on “Boudic” in the “Site related keywords” field. Unfortunately this positive result was then rather curtailed by the fact that very few of the sites actually had digital data available in an easily useable format. The situation was summarized by the curator in the following email extract “Out of the 39 sites, 28 appear to have digital data associated with them, although unfortunately none of it is currently available for download through the online catalogue. Much of the data was deposited before standards for digital deposition were issued and therefore it is in proprietary formats which require extensive processing before they can be made accessible. Many of the more recent sites have not yet been deposited with us, which means we don’t have access to any of the data which might be in better digital formats”. (pers com)

Despite a general overview story from London that “Excavations in the city suggest that few buildings, if any, had escaped the fire” (Perring p16), for the, admittedly limited, research purposes of this case study only data from the site BLE98 (Bush Lane House, 80 Cannon Street, EC4) was readily available in a tab-delimited format. However another one of the sites returned was No1 Poultry and data was also used from this very recent publication (Hill and Rowsome 2011).

For Verulamium the principle sources were similar to London, although there are nothing like the number of on-going excavations that there are in London. After a search on the Heritage Gateway returned 1884 records for the date range 60-62AD, the search narrowed down to the Verulamium museum. Again the responses to emails from staff were helpful, but it emerged that there were no excavations in recent years that had any digital data available. The main excavations with any Boudican period archaeology were carried out several decades ago and particularly by Mortimer Wheeler in the 1930’s. A useful suggestion was to look at the relatively recent publication synthesis of excavations at Verulamium (Niblett and Thompson 2005) which contains data in printed tables for the main excavations in the Verulamium area. The authors summarize the main evidence for activity in the Boudican era:

“There are still only seven sites known from the town where structures destroyed at about this time have been recorded with any reasonable certainty. These include timber buildings in insula XIV and XVII which were destroyed in the Claudio-Neronian period, and the insula XIX bathhouse which was extensively damaged at about the same time;” (Niblett and Thompson 2005, 150)

In view of this I did transpose the small amount of data on these excavations to a CSV spreadsheet and ran that through the STELLAR project tools (May et al, 2012) to create an RDF version of the data in a standalone triple store using Alegrograph Gruff software (Fig. 3).

The first settlement attacked by Boudican forces was Camulodunum. A search on the Heritage Gateway for the timespan 60-62AD returned 1510 records from the Colchester UAD. Unfortunately an enquiry to the Colchester Museum revealed that the Colchester UAD had recently had staff cuts so no member of staff was available to follow up questions about data in the Historic Environment Record. I therefore
directed enquiries to the main excavating organisation, the Colchester Archaeological Trust. I received a very informative and interesting reply including a personal eye-witness comparison “The Boudiccan destruction debris at Verulamium looks to have been much more modest by comparison, but London looks much more like Colchester” (pers com).

However no data was available for any further investigation on the Colchester evidence in time for the presentation or publication of this paper. A couple of recent excavation sites were mentioned, Head Street and Queen Street, but this really highlights one of the key issues with this line of research and broader semantic research questions in general. Because the more recent data is from sites which may still have a commercial archaeological interest, there is some resulting reluctance from archaeologists, prior to their own publication of data, to make that data available for re-use in research since other organisations may make commercial use of it.

In summary, the broad level searching on the Heritage Gateway for data for the period 60-62AD was able to identify a number of sites from London, Verulamium and Colchester, but there was no way to narrow down those searches to anything specifically related to Boudica, nor to geo-locate any of the London sites in an area related to St Pauls. It was possible to get a search based more accurately on the St Pauls area using the Greater London HER but it was not then possible to extract just the sites that had data relevant to the period 60-62AD.

Where sites were identified as being of relevance to the Boudican enquiries, largely by personal communication, it then proved very rare that any of those sites had records available in a digital format (CSV, or spreadsheet) that would enable transforming them to RDF or Linked Data. Most of the data from earlier (pre-2000) excavations remains ‘locked’ away, often as tables in printed publications. More recent data – from say the last decade - that is more likely to be in re-useable digital formats is harder to get hold of as it has often not been formally published and the information is still seen as having potential commercial value if nearby sites in urban centres are still likely to come up for re-development and future archaeological investigation.

3. Consideration of Broader Issues Raised by the Research Questions Case Study and Related Semantic Technologies Research

This section considers some of the information management implications there might be for trying to tackle new sorts of research questions using semantic web based technologies.

3.1 Potential for Semantic Technologies and Linked Data Based Research Questions

Where does this case study leave the potential for new research questions? Clearly the capability is still there to ask new, or perhaps more accurately, more complex, questions or examine new avenues of enquiry leading on from existing questions. But a major issue seems to be whether there is data that is accessible, or even available, in a format that can be used.
Single, fixed length university based research projects may provide ‘packages’ of project-specific data that can be usefully integrated and examined (e.g. looking at a range of projects around a study area over a number of seasons). But a real grand challenge for developing new questions would be to enable a cross-searching capability across all new interventions that take place under the development control process on appropriate national or regional scales.

What might be necessary to do this? A first requirement would be the ability to access an RDF/XML Linked Data record of all the investigations that are reported to National or Regional records. If resources such as the Historic Environment Records (HER) were available as Linked Data, even just at the level of a basic index record, then this would create the opportunity for other organisations to link further records to those higher level indices. In this way the HER Linked Data records could provide the ‘skeleton’ of a much richer, more fleshed out network of integrated research projects, reports and associated data for the Historic Environment. It could also provide opportunities for Local Societies and special interest or community groups to integrate their own local, specialist, or more personal records of the historic environment. This could include perhaps areas such as local history studies, artefact or scientific studies or historical document based research by relating such records to the Linked Data framework that could be derived from their local or regional HER.

An opportunity to do this might come from development of the OASIS system, so that it can output the index records of Investigations in RDF/XML and Linked Data formats. Currently the Heritage Gateway provides a portal across a proportion of the Historic Environment Records for England and it might be conceivable for the Gateway or another national resource to provide a Linked Data browsing functionality in the future.

3.2 New Research Questions Posed by STAR Outcomes Involving Spatio-Temporal and Stratigraphic Views of a CRM Matrix

The STAR project demonstrated, using archaeological extensions (Cripps et al. 2004) to the CIDOC CRM (Crofts et al.) that cross-searching data from previously un-connected archaeological data sets using semantic technologies is possible (Tudhope 2011). This in itself raises possibilities to ask new questions using data that has not previously been combined. One area that shows potential for exploring new research questions that can be interrogated using semantic technologies is the extension of the spatio-temporal relationships between different contexts recorded as part of the methodology known as Single Context Recording. In current archaeological work the main mechanism for expressing temporal relationships between single units of stratigraphy at present is the Harris stratigraphic matrix (Harris 1987). In practice the Harris matrix approach covers only two of the basic temporal logic relationships that exist between stratigraphic contexts: namely the Above/Below relationship (represented in CIDOC CRM terms by the p120 Allen operator relationship “Occurs Before/ Occurs After”) and the “Equals” relationship (CIDOC CRM P114 is Equal in time to). In practice an archaeological context can be deposited in time either before, or after, another context, or they can be deposited at the same time. These temporal relationships are fundamental to understanding the sequence of deposits on excavations and therefore currently receive the primary attention in the recording, and interpretive reconstruction of events. However archaeologists also already record a range of other spatial or temporal relationships. Examples of commonly recorded spatial relationships (often referred to as ‘Physical’ relationships) which are recorded as part of the archaeological record, include “fill of”, “filled by”, “cuts”, “cut by”, and in the case of built structures, “butts”, and “is butted by”. The STAR project was able to demonstrate the incorporation of the “fill of / filled by”
relationships into a stratigraphic browsing window (Fig. 5 – below), but more could be done to incorporate views of other relevant physical relationships along side stratigraphic ones, enabling more specific details of deposits to be interrogated in addition to the simple Above/Below relationship. This could enable researchers to develop and query on much richer patterns in the sequences of deposits, to find things like examples of layers that always occur in similar sequences in association with particular features such as hearths, burials, storage pits, etc.

Another aspect provided by a mapping to CIDOC CRM is the incorporation of the Allen temporal operators which identify other temporal relationships in addition to the basic two used by Harris matrices. These additional temporal relationships which may be recorded or possibly inferred include “Starts by” (which would equate to archaeological Terminus Post Quem); “Finishes by” (equates to Terminus Ante Quem); “Occurs during”; “Overlaps in time with”; and “Meets in time with”. These Temporal relationships are not currently used in the Harris matrix, but such temporal relationships, in particular the Terminus Post Quem and Terminus Ante Quem, do appear elsewhere in the archaeological record, perhaps most commonly in the attribution of spot dates to finds objects. The incorporation of these temporal relationships more closely with stratigraphic records could provide new opportunities for more detailed temporal searches. This would potentially open up opportunities for data mining to look for ‘patterns’ where archaeological stratigraphic relationships are combined with physical relationships in common sequences representing similar depositional events over common time spans.

The potential for combining more complex temporal relationships may also bear fruit in the ability to combine archaeological data with documentary records by extending the use of approaches that combine not just direct stratigraphic relationships, but also incorporate documentary temporal data (Holmen 2009).

In this way the CRM potentially provides a new way for constructing a matrix which is not simply a record of above/below or equals relationships, but rather a multi-dimensional view incorporating a range of conceptual entities that are inter-related by a variety of spatial, temporal, and spatio-temporal relationships. The challenge of course is then to be able to produce a suitable interface that can actually present that complexity of data inter-relationships. A prototype interface developed by the STAR project showed how presenting all the possible relationships between data contained in such a Conceptual Reference Matrix was possible but also demonstrated that such an interface could present considerable issues to the user in trying to navigate the information in meaningful ways (May 2008).

The first part of this paper has also sought to show that considerable work might be necessary to get a suitable quantity of data available to answer such research opportunities, although it should be noted that some of the more detailed stratigraphic based research might be perfectly applicable to internal organizational enquiries or across just a number of spatially related sites, rather than necessarily at a much greater regional or national scale.
3.3 Potential for Terminology Alignment – SKOS

One key factor in enabling semantic cross-search is having the capability to use standardized terminologies. One advantage of using the Historic Environment Records as the framework for any Linked Data initiatives is that they already do incorporate a number of the key standardized terminologies, such as the Thesaurus of Monument types and Periods in England. However, STAR also showed that there was considerable variation (sometimes even between projects within the same organisation) in the use of standard terms for data recording (Binding 2010).

Something that would be extremely valuable would be an online service for terminology alignment that could form part of the infrastructure for cross-referencing glossary terms in standalone systems to a broader SKOS constructed terminology service at a national level. Being able to map across the different terms would be a major step to advancing semantic search capabilities, at the level of removing the ambiguities brought about by such terms as posthole, post-hole, post hole, etc, but also to enable more sophisticated mappings between searches such as Boudica, Boudicca or even “Boudican period”, that could be used for connecting across a range of resources.

4. Conclusions and Future Directions of Research

The case study section of this paper began by considering some avenues of research enquiry that might be addressed by collecting data in suitable formats for conversion to semantic web and Linked Data formats. The intention was to consider and highlight the issues that the semantic technologies are most likely to be effective when addressing questions on a scale not so far readily addressed by current technologies. It seems one issue is that it will be hard to track down appropriate large-scale data sets because they will not be produced until the technologies are more mature for making use of them (comparison could be made to the position with large-scale spatial datasets before the advent of desk-top GIS).

The second section of this paper considered some of the broader issues arising from the case study and offers a number of possible requirements for data management in order to be able to adequately address questions of greater ‘semantic’ complexity, particularly if using Linked Data. The challenge to doing this can be that it does require data to be structured (and recorded) accordingly, else considerable, and perhaps unfeasible amounts of work would be necessary to prepare legacy datasets. Perhaps even more of a challenge of scale is that in order to achieve more significant criteria of ‘coverage’ and ‘completeness’ the data sets involved probably need to be more comprehensive and well structured and recorded to actually be appropriate for such use.

The possibilities are potentially greater with semantic technologies for answering, not necessarily totally ‘new questions’, but rather addressing new levels of complexity in questions involving more data, more inter-related data, with more complex relationships between different parts of the data or between different but related data sets.

The overall results of this particular case study suggest that currently, at least for certain avenues of enquiry, there is limited availability of data on a suitable scale and in suitable formats to provide answers to more complex cross-project, and particularly cross-organisational research questions. Perhaps understandably, currently most data is held in a way that serves intra-project and intra-organisational interrogation. The actual amount of RDF/XML data produced from the case study was, disappointingly, minimal and consisted of event records for excavations in London and Verulamium that had either some documented notes or context level data relating...
to the Boudican period. These outputs are not presented as completed research but merely as indicative of the limited data available. The data outputted that was produced in RDF/XML is summarized by the Gruff diagram shown in Figure 4. The somewhat anti-climactic nature of these results may serve to emphasise the point that the potential of semantic technologies to answer new questions rather depends upon being able to amass significant quantities of new or existing data and integrate and query them in ways not previously possible.

On a slightly more positive note, perhaps it is understandable to a degree that we have not been providing archaeological data in such ways to date because there have not been readily available technologies and tools for using the information on such a level — although some resources may provide exceptions to this, it still appears to be very rarely the case (Richards p106). But as more and more data is created in digital formats then the possibility, and many would argue the requirement (Finch 2012), increases for that data to be openly available (perhaps if not immediately then after a generally acceptable amount of time) for re-use and research by others. This ‘open’ approach to research data from sources with public funding may raise challenges for parts of the archaeological sector for which detailed data may carry some commercial advantage.

A further possibility might be for records of investigation events to be made available as Linked Data metadata records to a level of detail which would at least allow cross-search between investigated events relating to similar archaeological research themes such as site types and periods and incorporating key artefacts, ecofacts and similar levels of information. Such a configuration of linked metadata could then become the framework upon which more detailed records of individual projects could be made accessible. In the UK, such a framework of event records does to a degree already exist in the form of regional HERs along with national overviews of such records such as the Heritage Gateway and the OASIS reporting system. Extension of such resources to provide semantic and Linked Data interoperability functionality could be a worthwhile next step, if enough of the sector were prepared to provide their own data in compatible and open RDF/XML formats, and even more ideally to publish their data themselves as Linked Data.

References


Websites:


Connecting Archaeology and Architecture in Europeana: the Iberian Digital Collections

Ana Martínez Carrillo, Arturo Ruiz and Alberto Sánchez

Universidad de Jaén, Spain

Abstract:
Recent years have seen the development of new communication channels for spreading and disseminating cultural heritage through the Internet. In 2005 a new project for building a digital library of European culture was funded, known as Europeana. The aggregation of content in this network is carried out gradually through direct cooperation between institutions and Europeana or through the foundation of projects developed for this purpose. In this sphere of action the European Commission has funded the project CARARE (Connecting Archaeology and Architecture in Europeana, ICT Policy Support Programme 2009, c. 250 445). CARARE aims to increase the quantity and quality of digital content in the field of archaeology and architecture available in Europeana. It also provides aggregation services for users and enables access to 3D and Virtual Reality content. The Andalusian Center of Iberian Archaeology is member of CARARE as content provider. This contribution focuses on the aggregation of 2D and 3D contents relating to the archaeological remains of the Iberian culture.

Keywords:
Iberian Digital Collections, Metadata Exchange, Aggregation of 2D and 3D Contents

1. Introduction

Recent advances in the development of Internet technology, its easy accessibility and the increasing potential it has to be a useful tool in various aspects of life are gradually changing the form, content and the direction of archaeological research. In this sense, excavation methods and data collected and published concerning Cultural Heritage should be reorganized, taking into account the new possibilities for displaying and sharing information.

One of the initiatives to make accessible and to disseminate cultural content through this channel is represented by Europeana http://www.europeana.eu/portal/. This initiative dates from 2005 and its main objective is to make available over the Internet material related to European Heritage. Through this portal are available resources and digital collections from museums, libraries, archives and audiovisual archives in Europe. Currently this site has more than 15 million items which include images (drawings, maps and photographs), texts (books, newspapers, letters, diaries and archival documents), audio (music and radio) and videos (movies and TV).

Around 1,500 institutions are contributing to the development of Europeana, including the British Library in London, the Rijksmuseum in Amsterdam and the Louvre in Paris. Through these different contributions the history of Europe from Prehistory to the Modern and Contemporary age can be explored.

Architectural monuments and archaeological sites are an essential part of the tangible European heritage. Following the initiative of establishing the digital library of Europeana as a single point of access to the European cultural heritage network, the Best Practice Network of CARARE was created in 2009 to increase the quality and quantity of contents related to archaeological and architectural heritage available to Europeana users (Hansen and Fernie, 2010).

The main objective of CARARE is to integrate 2 million digital resources; of which approximately 1 million are unique architectural
monuments, landscapes and artifacts. The content belongs to a large and diverse group of organizations and European institutions, with equally diverse cultural objects originating from Prehistoric and Iron Age archaeological sites, archaeological sites from the classical world and historic buildings. These digital resources are also heterogeneous (paintings, prints, photographs, archaeological and architectural maps, sections and drawings). In addition, 3D representations that have been developed in recent years will be incorporated.

Initiatives have recently been developed to provide metadata and links to digital resources published on the internet for their aggregation into Europeana, based on correspondence through certain rules of the original data with the metadata proposed by the semantic model of Europeana (ESE). (Europeana v.01). Following this model, the integration into CARARE is based on the creation of an intermediate repository that ensures the integrity, authentication and enrichment of metadata coming from diverse and heterogeneous European collections.

2. The CARARE Scheme

The semantic model proposed by CARARE is based on existing standards. These are: CIDOC Archaeological Sites Index Core Data to Historic Buildings and Monuments of the Architectural Heritage, CIDOC CRM, MIDAS Heritage, POLIS DTD, LIDO, GIS (Geographic Information System) metadata and ISO 8601.

The most influential standards in the construction of the CARARE scheme have been LIDO and MIDAS. LIDO (Lightweight Information Describing Objects) is a standard based on the working group for the aggregation and exchange of information of CIDOC (ICOM / CIDOC) and is the result of the efforts of CDWA Lite, the Museumdat working group, the Spectrum community, the Documentation Committee of the German Museum and the ATHENA project. LIDO aims to collect information on museum objects and to add content using different standards and schemes developed in Europe and elsewhere (CDWA Lite, CIDOC CRM, Museumdat and SPECTRUM). This scheme covers seven areas for each record of an object, including in each record metadata of identification, category, description and administrative information.

On the other hand, MIDAS (English Heritage) is a standard which indicates the type of information to be collected for correct information exchange and allows the conservation of long-term knowledge of historical surroundings.

MIDAS has a structure consisting of three levels: Threads, the highest level of information; information groups, which indicate the specific standard to be included in the description of a particular record, and finally the information units which are the basic elements that comprise an entry.

The main findings of a comparative study between the LIDO and MIDA schemes show that MIDAS is a more general standard which covers several areas of cultural heritage management, while LIDO focuses mainly on the description of museum objects, providing rich semantic information for information exchange.

In general, the two schemes have different orientations and serve different purposes and functions.

The CARARE scheme has been constructed from a detailed comparative analysis of both schemes. It includes a language for describing immovable monuments, focusing equally on the objects and their digital representations and thereby covering a domain that falls within the scope of this project (Papatheodorou et al., 2011).
The CARARE scheme is designed primarily to allow the aggregation of data into Europeana. These contents, as well as being largely heterogeneous both at documentary level and for the different digital resources referred to, must be integrated at the same time following the format used in Europeana, which has been defined as EDM (Europeana Data Model) and is currently under development. The EDM is the latest proposal for structuring the data integrated, managed and published in Europeana. The main objective for the adoption of this model is to enable users to find and insert content into the Europeana Semantic Web. The main advantage of the Europeana Data Model is that it is not subject to any standard used by a specific community, but rather is developed within the framework of the semantic web that can be adapted to the different ranges of standards used to date (Chambers and Schallier, 2010:116).

Conceptually, the CARARE scheme is divided into four themes that comprise the top level:

- **Identification of the cultural asset**: This includes descriptive information and metadata about the monument, historical building, archaeological site, the artifact / ecofact, etc.

- **Digital resources**: digital representations and sources of information relating to a particular cultural object (images, text, video, audio, 3D models).

- **Activity**: these are events that the heritage asset has taken part in, such as Creation, Field investigation, Research and Analysis of Historical events, etc.

- **Collection**: this is a collection level description of the data being provided to the service environment.

For each of the four themes of higher outline level CARARE defined a comprehensive set of descriptive and administrative metadata:

- **Identification**: name (or title) and identification number.

- **Description**: Refers to a textual description of the characteristics of cultural resources.

- **Spatial Information**: Evidence from names of places, geo-spatial coordinates and other spatial data.

- **Temporal information**: absolute dates or other information relating to the time span.

- **Actors**: data from people, organizations or groups and their roles.

- **Rights**: associated with the object or metadata

- **Relationships**: between top-level items.

- **Registry information**: metadata about the record.

In addition to the global elements for the definition of a digital resource, the following additional information has been included:

- **Format**: The format of the resource file, for example, MIMES.

- **Theme**

- **Publication status**: name of publisher, place and date of publication

- **Type**: nature or genre of the resource.

- **Link**: the URL where the resource is published digitally.

- **IsShownAt**: a link to the web page that contains the digital object and contextual information.
The location of the resource metadata refers to web addresses where additional information on the digital resource can be found.

- Rights: the rights associated with the digital object (copyright, access rights, copyrights, etc.).

3. Implementation of the CAAI Collections and the CARARE Scheme

The Andalusian Center of Iberian Archaeology (CAAI) is involved in the CARARE project as a content provider. The collections that will be available in Europeana are the following:

- Image databank for the teaching and dissemination of Iberian culture. This collection consists of 2,000 cultural resources that illustrate and explain various facets of the Iberian world. This collection includes content relating to objects held in museums or state institutions as well as information regarding the populations who occupied the Iberian area.

The 2D images include drawings of plans and profiles of archaeological excavations, archaeological drawings, aerial photographs of settlements (Fig. 1), photographs of materials and images of archaeological methodology. The 3D models incorporate the latest interventions performed by CAAI in the cemetery of Piquías (Arjona, Jaen) and some metallic elements documented in recent surveys carried out at site of the battle of Baecula (Santo Tome, Jaén). From this collection only the images have been conserved, without any associated information. For this reason, once the metadata schema of CARARE was defined, an on line repository was developed with the required fields reflected in the CARARE scheme.

- CATA-Collection (Archaeological wheel ceramics in Andalucía). This reference collection makes available on-line information about complete vessels and pottery fragments. In addition, registered users can expand the reference collection with new materials.

Project information and access to the application are available at the following web link: http://cata.cica.es/. The metadata for the CATA reference...
collection were selected according to certain basic protocols for the standardization and integration of information of archaeological pottery. These protocols have been developed in successive meetings between the different members of the project, a consensus being reached among researchers of Iberian, Roman and Medieval archaeology.

This being an on-line collection, the integration of metadata from both collections (Fig. 2) has been performed, indicating the equivalent metadata of both collections. The CATA collection was originally divided into the following levels:

- Metric variables: these are taken based on an image which contains a drawing of the ceramic vessel. The measures have been classified as:
  - Basic measures: diameter, height, volume and weight of a complete vessel.
  - Complementary measures: allow the definition and specific numeration of the most significant parts of the morphology of a vessel such as the rim, the handle or the base.
  - Qualitative variables. These refer to the production process of the vessel. So aspects related to the modeling of the vessel, the type of cooking, type of clay and additive elements are included. Also a description of the morphology of the vessel (rims, handles and base) and a description of surface treatments and decorations have been included in this section. A form relating to chemical analysis of the vessel can be completed in this part.
  - Conservation variables. These variables concern the condition of the pottery vessel (complete or fragmented), possible alterations and the treatment that the piece has received or may be proposed for its restoration and preservation.

- Contextual variables. Finally, a section relating to the context in which the ceramic vessel has been documented has been included. A vessel is located in a specific temporal and spatial context. The identification of these contexts allows dating, functionality and distribution. The spatial information has been included from geographic coordinates.

Another objective of the CARARE project is to integrate 3D models and virtual reality representations of historical monuments, archaeological sites and museum objects into Europeana for touristic, educational, researching and dissemination purposes.

A 3D digitization of all the models to be integrated has been carried out in a very diverse manner, including 3D scanning techniques using a laser scanner, image modelling (Nilsson 2007) and other techniques. Each of these techniques has specific requirements. A homogenization of the different 3D formats should be carried out in accordance with the specifications of Europeana. This requires an open and standardized format and means that in practice the chosen format type can be used by the largest number of users. At this time there is only one technology that coincides with almost of these specifications, the PDF format. Since 2005 (Acrobat 7) the PDF format has been expanding.
its 3D capabilities (Fig. 3). The current version is the Acrobat 10 and has been in use for 6 years (Pletinckx, 2011).

In the case of the CAAI collections, 3D digital resources will be integrated from both collections. The conversion into 3D PDF has been carried out using the exchange file U3D. 3D PDF documents are being integrated into the local repository as digital resources for their later incorporation into Europeana (Fig. 3).

4. Mapping

The structure of CARARE adapts best to an on-line repository with the metadata established in the latest version of the schema (v.6.0.1). This repository is hosted by Project CATA (http://cata.cica.es/carare/index.php/acceso) and is based on LAMP (Linux, Apache, Mysql, PHP) applications, which are efficient and low cost.

The records of the collections have been inserted in two different ways:

- The data of the 3D models of the CATA collection were migrated to the local repository. The correspondence was established between the metadata available in CATA and the metadata of the CARARE scheme.

- As to the collection Image databank for the teaching and dissemination of Iberian culture, the information linked to the 2D and 3D models was scarce, unnecessary or non-existent. The metadata of the heritage assets and digital resources therefore had to be inserted from the start for the 2D and the 3D models. This allowed further data selection and deletion of irrelevant or redundant information for Europeana.

The local repository allows data addition and edition of the heritage assets and digital resources, as well as insertion of archives with information in the following formats: image files (.TIFF, .JPG), video files (.FLV), and 3D models in .PDF format.

The information was first ordered and structured following the guidelines of the CARARE schema, then uploaded to the CARARE repository and finally validated with an online mapping and validation tool (MINT) (Fig. 4). In the case of the CAAI collections, the information was compiled in an XML file and then uploaded to the CARARE repository following the parameters defined in the latest update of the schema.

Several tests have been carried out on the uploaded XML files to confirm that the data schema is right. The latest XML file uploaded, which contained 3D models as PDF files, was validated and is ready for publication in Europeana.

5. Conclusions

First of note is the opportunity for on-line dissemination of all the elements included in the Europeana portal, special interest being given to the diffusion of Iberian culture through the integration of the collections described above. Furthermore, this integration also allows the updating and adaptation of the metadata information stored in local repositories to existing metadata schemes and the opportunity to standardize the content relating to European cultural heritage.
Within the framework of project development, it should be noted that the architecture of the CARARE scheme is represented by a central repository that stores all the metadata before being assigned to Europeana. The concentration of all metadata objects in a single system creates added value by allowing content providers to enrich their objects through semantic relationships. The use of a central system also facilitates the preservation of data from different collections.

Finally we must also highlight the inclusion of 3D digital resources, which will provide added value to the Europeana portal. 3D visualization enables the enhancing of animation capabilities and facilitates the understanding and exploration of the content, making it subsequently more attractive.

Acknowledgements

This work has been made possible by funds received from CARARE project (Connecting Archaeology and Architecture in Europeana, ICT Policy Support Programme 2009, c.250,445), the Andalusian Center of Iberian Archaeology (CAAI) and FEDER funds.

References


**Open Access Journals in Archaeology and OpenAccessArchaeology.org**

**Doug Rocks-Macqueen**  
*University of Edinburgh, UK*

**Abstract:**  
This paper will briefly explain the current issues facing Open Access publishing in archaeology and some of the current solutions. This examination begins with a brief look at what Open Access is and how it relates to other fields of archaeology like Open Data. This will be followed by an investigation of, but not all, the problems facing Open Access in archaeology. Next, some of the solutions being implemented at the website OpenAccessArchaeology.org will be surveyed. Finally, some of the plans for future improvements will be discussed.

**Keywords:**  
Open Access, Publishing, Archaeology, Open Data

1. **Introduction**

There has been a flurry of media events regarding Open Access (OA) publication at the end of 2011 and into 2012. Some of this has been related to national policies such as the United Kingdom’s government looking into making all of the research they fund OA (Willetts 2012). Other events have been on a more localised scale such as the criticism, and subsequent change in OA policy by scholarly societies such as the American Anthropology Association (Report 2012). Archaeology has seen its fair share of OA news with Archaeological Institute of America coming out against OA (Bartman 2012), and subsequently reversing its stance due to pressure from the archaeology community (Bartman 2012).

However, for all of this exposure to OA there is very little research into OA and its relationship to archaeology. Published articles on the subject of OA in archaeology-based publications are practically non-existent, a search of any repository or search engine will find only fleeting references to OA in an archaeology context. Thus, this paper will serve as a first to introduce the archaeology community to OA and some of the problems it faces within archaeology. The purpose of which is to start a conversation about what needs to occur for archaeology to begin to fully benefit from OA.

2. **What is Open Access**

The Budapest Open Access Initiative describes Open Access as:

“There are many degrees and kinds of wider and easier access to this literature. By ‘open access’ to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.” (Initiative 2002).

There is a wide range of ways to implement this concept with some authors or publishers limiting reuse to non-commercial applications to control the ‘integrity’ of their work. In some cases, publications are not in easily accessible
formats for others to index or pass as data to software which makes it hard for others to use. However, all OA publications follow the same basic idea, free access to knowledge.

2.1 When Free is not Free

It is important to point out that the free access in Open Access only applies to viewers accessing the information and does not mean that OA is completely cost free. There are many assorted costs from website hosting, to paying for copy editing of papers. Sometimes these costs are covered by an organization. In other instances volunteer donations help keep Open Access publications accessible, the Arqueologia Publica journal is an example of such a publication. In other instances, it is the author who pays the costs of publication which is a model used by Ubiquity Press for their archaeology journals.

Moreover, it is important to make the distinction between legitimate avenues of publication and those that are trying to take advantage of authors. There are some less than scrupulous publishers that use the author pays model of OA to scam potential authors. This involves accepting money to publish papers but not undertaking any peer review. There are lists of such publishers to avoid which can be found on the internet e.g. Beall’s List of Predatory Open Access Publishers. Unfortunately, their actions have hurt the image of OA and it is important to distinguish OA from vanity presses that do not undertake peer reviews.

2.2 Open Access, Open Data, and Grey Literature

OA is very similar to the concept of Open Data, sharing the same goal of free access to information. The difference is that OA is applied to publications that analyse data while Open Data is the rough un-analysed data. However, these lines are becoming ever more blurred with new projects such as the Journal of Open Archaeology Data. In the future the two issues may be indistinguishable but at the moment a separation is still maintained.

Within the remit of Open Access, publications can be roughly divided into two groups; Grey Literature and journals. Grey Literature for OA in Archaeology is defined as site and project reports produced for the needs of Cultural Resource Management. While journals are the result of independent study, sometimes peer reviewed but not always, that are published as speciality periodical publications. There is overlap between these two areas with many heritage management projects being written up as journal articles. OA journals make up a very small percentage of the ‘Open’ ecosystem (Fig. 1) and they will be the focus of this paper.

3. Open Access and Archaeology

Open Access and archaeology, like other disciplines, has a very recent association. It is only possible through the lower transaction and distribution costs provided by digital publication e.g. lower cost to download a
digital copy of a paper then it is to print and mail the same copy. At the current moment no publication takes credit as being the first OA archaeology publication. Internet Archaeology, which was first published in 1996, originally started out as an OA publication, but is now a hybrid OA publication (Xia 2006), indicating that OA and archaeology have at least been connected for 15 years. However, in that short time OA has greatly expanded. Figure 2 demonstrates the massive growth seen in Journal Publishing within archaeology in the last few decades.

These numbers are not perfect as they are limited in scope to journals that at least some of their articles in English. Furthermore, it is limited to journals that publish articles primarily concerned with archaeology issues and excludes a whole range of journals that will publish archaeology related articles along side a whole host of other subjects e.g. history, classics, anthropology, etc. Still there is a clear example of an increase number of archaeology journals, specifically OA journals and closed Commercial Publishers.

4. Problems

In early 2011 research was undertaken to find as many OA journals as possible to serve as a dataset to look at the problems that affect OA publications. Many of the journals were found through the list kept by Chuck Jones on the blog Ancient World Online (Ancientworldonline.blogspot.com/); while others were found by conducting searches of journals databases e.g. Directory of Open Access Journals, Science Direct, etc. The loose definition used to define Open Access Journals was any journal that had a possible of making whole issues, as opposed to only certain individual articles, Open Access. However, publications that did have a rolling wall policy, where the journal issues were not open access until after a set period of time, were included in that rolling wall policy was five years or less.

In total, over 270 periodical publications were found and recorded, out of these roughly 217 are still publishing. The full list of publications, are currently stored on the website Open Access Archaeology which is discussed in more detail below. At the moment this is a rough estimate as the exact number is not known due to irregular publication schedules masking when a publication has ceased publishing or simply gone on hiatus. This dataset served as the bases for the following examination of issues surrounding OA publishing in archaeology. It is not a complete list of all OA journals as new ones are started every month. Moreover, the following list is not a complete list of all
possible problems. It is mainly an examination of some of the more pressing problems found in this dataset.

The recent growth in OA publishing has led to fragmentation in OA publication within archaeology. The 270 publications examined are spread out over 300 websites. Many times of OA publication will have multiple websites for different purposes. Some will have their most recent publication on the Journals website but then have back issues archived on a separate website. Some with have an older website and more current one but have elected to keep the older website up so as to keep the URL’s working. This leads to confusion among readers about which websites are the most current ones or if a journal is still publishing. When one lands on a webpage that was last updated 10 years ago the there is the tendency to believe that a journal has ceased publishing.

Almost all of the publications are run by organizations, groups, or individuals. This results in almost no centralization of information on OA publications. There are a few exceptions such as Ubiquity Press, ADS, Hrcak, RAE, CISC, and Persee websites which host more than a single journal but none of these host more than a dozen OA Archaeology journals. This means in order to discover OA publications a person would have to visit dozens to hundreds of different websites. Surprisingly, there is no central list of OA journals for archaeology, the Directory of Open Access Journals contains only several dozen and AWOL blog focusing only the ancient world and including history and classics publications.

As a result of the fragmentation there is very little standardization among OA websites. For example, some are created using Flash while others use standard HTML. Those websites that use Flash or Flash plugins are not currently searchable by most search engines. Not being listed in search engines can almost kill most websites organic growth as no one can find their websites through search engines.

However, several journals have adopted the use of Open Journal software to manage their websites, which is starting to elevate this problem somewhat.

While some websites use Open Journal software most do not, which results in a huge variation in quality among sites. A little less than two-thirds of the journal websites have any sort of search capabilities. This leaves a user with the option to either include the journals name in search results on a search engine on Bing, Google, etc or to open every document or webpage on the sites in hopes of finding the information. Furthermore, for journal websites, that are still publishing, only 45 out 217 of them have RSS feeds. Most OA websites have almost no way to let interested parties know when new publications are published.

Even other methods of informing potential viewers of material, such as social media, are almost non-existent. Less than 10 journals having linked Facebook pages to their websites and those that do are mainly published by Ubiquity Press. Twitter is even less represented on OA Journal websites. While no evidence of other social media outlets such as Youtube or Blogging was found. It would appear that most OA journals have no way or intention of expanding their audience through alternative channels of communication.

Finally, the last problem encountered was not with websites but with archaeologists. When emails were sent out to archaeology list serves dozens of responses would come back with essentially the same question, “What is Open Access?”. Even the recent anti-OA statement put out by the president of the AIA did not correctly identify OA; they confused public lectures with OA. While not every archaeologist is unaware of what OA is or that OA publications exist in archaeology, there seems to be a significant body of archaeologists who do fall under those categories. There is a definite need to educate a percentage of the archaeology community about OA.
5. Solutions

The website OpenAccessArchaeology.org was created with features to solve these problems, along with several social media related annexes. The website is run by a group of volunteers using the free website creation and hosting service Weebly. Using Weebly means that there are no hosting and creation costs other than paying an annual fee for a domain name, currently less than ten dollars annually. This allows the website to be independently run and operated without the need for institutional support.

To address the problem of discovering OA publications the website hosts a searchable database of journals created with Google Fusion Tables and Google Charts. One can search for journals based on pre-defined categories such as: Peer reviewed or not; geographical speciality, if any; the topic(s) covered; if it is still publishing or not; a journal's language. While it is possible that many OA archaeology journals are missing from the database at the current moment it is the most comprehensive list of such journals on the internet.

The issue of finding specific information in publications is addressed through the use of a custom search engine on the site, Open Archaeology Search (OAS). OAS is a Custom Google Search engine designed to search the webpages of the OA journals in the directory, and only those webpages. It also targets sites like Academia so that many other OA resources regarding archaeology are included in the search. The result of which is the ability of a person to search through hundreds of OA journals without needing to look at each individual PDF or webpage.

Letting users know when a new publication has been put online or that OA resources are available has resulted in a more labour intensive undertaking. OpenAccessArchaeology.org runs a Facebook page, Twitter account and Tumblr blog to inform people of new publications. While the postings on these social media sites are primarily aimed at letting users know about new articles some are aimed at letting people know about the general OA options available in archaeology. While not a “viral” sensation, at the moment, the combined social media channels has managed to attract over a 1000 followers/likes/subscribers in less than 6 months.

Social media also serves a second purpose which is to educate others about Open Access. While many of those who followed the different social media outlets at first were already interested parties there has been a steady stream of new users interacting with the different social media outlets. These are people who typically have no idea what OA is and what it has to do with archaeology. The result of which is a general expansion of the social media followers/readers/friends from archaeologists known to be involved in OA to those who have not visible connection to the concept.

6. Discussion

These solutions treat the symptoms of the problem but not the underlying causes, even then there are limits their effectiveness. As mentioned there is great variability in how websites are constructed. For those websites that rely on Adobe Flash it is not currently possible to index and search them using the Open Arch search engine. The use of social media takes up a large amount of time. While it is possible to mechanise some of the steps for posting information it is not possible to eliminate all requirements for human work. This is a drain on volunteers’ time and there is not always someone available to cover every possible need.

Another drawback to these solutions is keeping the information up to date and making sure that URLs still work. With 300 plus websites being tracked it is not uncommon for webpages to be moved or even websites to be discontinued. Several OA journals
have gone offline since first being recorded: Elektronski Casopis, Anarheologia; Journal of Interdisciplinary Studies in History and Archaeology; Mono y conejo: The journal of the Mesoamerican Archaeological Research Laboratory. In some cases these journals have hard copy versions of their issues but an increasing number of OA archaeology journals are born digital. When these websites cease operating there publications may not have any backups.

Considering that OA is a relatively new aspect of archaeology publications it is not surprising that there is a lack of standards and large variability in quality. Ideally, OA publications in archaeology will begin to move towards standardization, either through the coalescing of varies projects into several large repositories or projects or the adoption of universal standards. Already ADS in the United Kingdom hosts the back issues of several journals in OA form. Other data repositories such as tDAR or Open Context are only just getting starting out and while they concentrate on Open Data as already discussed the two issues are very close. It may be that these larger data storage organisations take a more prominent role in the OA movement within archaeology.

Conversely it may be that the adoption of universal or at least OA archaeology wide standards greatly improves some of the issues raised. Already several journals use the Open Journals management to run their websites. A wider use of such a system would allow for greater ease of indexing websites for search engines. The system also allows for internal searching of the website and for the creation of RSS feeds either RSS 1.0 RSS 2.0 or ATOM. This system is still being updated which means it will be able to adapt to changing web standards. However the use of these features in a single large repository would also accomplish the same results.

Regardless of how the results are achieved the important part is that they happen. When this paper was first presented and mention was made of Journals that had cease publishing and whose websites had gone off line one of the audience members asked, “What about OpenAccessArchaeology.org what happens when the website goes offline?”. The response of course is that is the hope. Once the issues facing OA are resolved there will be no need for the bandage approach that the website serves. Hopefully, this will occur sooner rather than later.

References


Bartman, E. 2012. From the President: Open Access. Archaeology 65 (3).


Willetts, D. 2012. “Open, free access to academic research? This will be a seismic shift.” The Guardian, 1/5/2012.

SVG Pottery: Upgrading Pottery Publications to the Web Age

Stefano Costa
Università degli Studi di Siena, Italy

Abstract:
Line drawings of archaeological pottery are normally published and shared in traditional ways, even in digital form. Raster or opaque vector images are lossy and do not convey to readers or users all the information the original line drawings contain. Is it feasible and sustainable to share and publish line drawings of archaeological pottery on the Web, retaining all the rich information? SVG is a native Web format for vector images and is therefore the best choice, but it does not come without problems.

Keywords: Archaeological Illustration, Pottery, SVG, Vector Graphics, Web Publishing

1. Pottery Drawings and their Digitisation

Pottery line drawings are a very specific and very large subset of archaeological data. They are made according to a variety of standards (e.g. Adkins and Adkins 1989, pp. 164-170; Collett 2008), but their common feature is to include a geometric representation of the vessel, often around an axis of symmetry. In digital terms, a geometric drawing is vector graphics. As a consequence, pottery drawings are, implicitly, vector graphics. However, even with digital publishing workflows, it is common for such drawings to be embedded as static, raster-only images, losing all the machine-readable information that is available in the vector image resulting from digitisation.

In the following sections, a manual process of both drawing and digitising (tracing) is assumed. The main reason for such an assumption is that the vast majority of line drawings are subject to a manual process, and it is realistic to target the existing conditions. However, automated tracing of hand-made drawings or other means of obtaining a digital representation of a vessel (e.g. Melero et al. 2004; Kampel, Mara and Sablatnig 2006) result in similar intermediate outputs, and can therefore be approached in the same way.

2. Other Formats: Why Not ... ?

There are several possible digital formats for pottery drawings. Some are commonly used, some are not, but all have in common one thing: they are almost never used for publication (with the exception of PDF), where print or raster images are the current standard. An example of raster images used to disseminate line drawings of archaeological pottery is Roman Amphorae: a digital resource (University of Southampton 2005), where a typological collection is presented and for every type of amphora one or more profiles are available as JPEG images.

In this section we discuss why most formats are not suitable for Web publishing.

DWG is a proprietary format developed by Autodesk Inc., and it is the native format of the AutoCAD family of software programs. AutoCAD is often used for digitising pottery drawings, for several reasons. First, it is used for other common tasks of archaeological routine, such as site plans. Second, it has a comprehensive set of drawing tools. Third, AutoCAD guarantees good-quality printed
graphics. But DWG is not a good choice. The first main problem with DWG is that most often it is used, but not chosen. Rather, it just happens to be the default format of a popular software. The second problem is that DWG is proprietary. Every version of AutoCAD introduces a new version of the format, incompatible with earlier versions. There is no open source software capable of reading and writing all DWG versions. No web browser has or is likely to have in the future support for DWG. As a consequence, DWG is not suitable for being published on the Web.

DXF is an interoperable version of DWG. Its specification is open (even though it remains property of Autodesk Inc.) and there are several open source libraries and programs capable of reading and writing DXF. However, DXF files tend to be very large in size compared to the corresponding DWG files, and have the same complex framework of incompatible versions.

PDF is an open standard, registered with ISO and part of lots of recommendations as format for archiving documents of many different types. PDF is a vector format, that is, vector graphics features (lines, polygons, text) can be scaled retaining their quality. However, objects contained in PDF are almost completely unstructured: if three lines form a triangle, this is not relevant in PDF, and therefore there is no way to extract the information “this file contains a triangle”. For this kind of purpose, pottery drawings are no different than triangles. PDF is suitable for archiving because it is easy to create but generally not easy to edit and PDF documents are usually intended to stay as they are. Images can be included in PDF also as raster images (with JPEG compression), so using PDF for vector drawings is an ambiguous recommendation. Moreover, PDF on the Web is an ambiguous presence: despite being very widespread, it is not natively supported by web browsers (unless using additional plug-ins) and it is treated as a downloadable file. PDF files containing pottery drawings are for example the ones published by the FOLD&R journal.

EPS and AI are two variants developed by Adobe Systems Inc. for their Illustrator software. They share almost all problems of PDF, and they are not native Web formats. To make things more complicated, recent versions of AI are proprietary and not interoperable.

GML is an XML language and an open standard of the Open Geospatial Consortium. GML can be used to describe any spatial feature, as commonly seen in GIS. Because all vector spatial features are also geometric features, it is possible to represent a vector graphics object in GML. GML has two significant limitations. First, it is not a vector graphics format, and normal image viewing programs are not able to deal with it. Second, it is deeply rooted in the geospatial domain, and therefore all data is assumed to be in a geospatial coordinate reference system (CRS). If no CRS is specified, WGS84 geographic coordinates are assumed, and most importantly there is no straightforward way of marking a GML file as having a local, non-geographic coordinate system.

3. SVG

SVG is a well-known standard W3C format for vector graphics on the Web, developed by W3C since the late 1990s, and has now major support from almost all browsers, including mobile platforms, whose importance cannot be overstated. SVG is an XML language, SVG has rich graphical capabilities, but it’s not without issues for representing real objects (or pieces of them). The main problem when choosing something other than SVG is that publishing and disseminating on the Web is difficult, if not impossible.

There are several programs that have native support for editing and exporting to SVG. A comprehensive catalogue of all these programs is not even attempted here, and we will concentrate on two of them, namely Inkscape and Adobe Illustrator. These two
programs are not necessarily the best available, but they show a combination of adherence to standards, widespread usage, user-friendly interface and, in the case of Inkscape, it is free and open source and cross-platform. In any case, it should not be assumed that SVG saved by one is the mentioned programs is perfect, nor that exporting from the graphics software should be the last step in the process.

In fact, the advantage of SVG being an XML language is that there are countless tools and software libraries available for automated processing, extraction of data, customised representations and converting from vector to raster formats. Therefore, SVG files can be processed in a number of ways, and the choice of this format leaves all possibilities open. As an example, to facilitate the maintenance of a catalogue, a simple non-interactive program was written in the Go programming language by this author, that automatically extracts metadata from all SVG files in a given directory, and creates an HTML index with previews, titles and descriptions of drawings. This program is still rudimentary, but it is already available under the X11/MIT license and when compiled runs on any operating system (CataloGo).

4. SVG in Archaeology: State of the Art

SVG is already used in archaeology, but not widely. One possible reason is that in archaeological vector graphics there is a focus on software rather than documents, and that poses several problems for data archiving and preservation, as already pointed out above with regard to DWG.

However, SVG is the only format for vector graphics recognised as “suitable for preservation” in the Archaeology Data Service / Digital Antiquity Guides to Good Practice (Niven 2009). On the other hand, the same Guides contain separate recommendations for CAD-based data, suggesting DXF as a preservation format. CAD can be classified as a subset of vector graphics (depending on the definitions, but also on practical grounds), leading to a potentially incompatible set of recommendations.

SVG seems to be used mostly in the geospatial domain (to the best of our knowledge, the only exception is Pagi 2010), that is, for representing features commonly managed in GIS, such as site plans (Charno 2007; Robinson, Campbell, Hodgkinson 2010) or sections (Wright 2006). The gap between geospatial archaeology and other branches of the discipline in technological innovation is beyond the scope of this paper, but acknowledging it is important, because it affects how far an experiment can be fetched, as in the case of SVG for pottery drawings.

There is one important case of SVG being used for pottery drawings: the Greek, Roman and Byzantine Pottery at Ilion (GRBP) digital publication (Heath and Tekkök 2009). In that case, SVG is used as the original archiving format for the hundreds of drawings, but it is not exposed in the web pages, where a “safer” JPEG version is used. GRBP represents a significant source of inspiration for the work presented here.

5. SVG for Pottery Drawings: the Good and the Bad

Among the reasons why SVG is important in the specific case of pottery drawings, some are significant for the purpose of sustainability. First of all, SVG can be used for disseminating high-quality vector graphics on the Web instead of raster images: in other words, this means that drawings will have a quality equal or better than a 1:1 print, and it will be also possible to perform manual or automated editing (e.g. based on personal taste, or requirements from publishers, colours and hatchings can be easily changed accordingly). Secondly, SVG works across most vector graphics programs, and has a rich community focused around the format...
and not the software. This gives a higher sustainability to archives of SVG drawings: should one program become obsolete or unmaintained, the migration path would be comparatively easy, because no conversion of format would be needed. Finally, dissemination on the Web is a good way to preserve data, thanks to the existing mirroring services, and to the inevitable downloads that will happen.

SVG is not without problems though, and it should be clear that SVG is far from perfect as a native format to create and archive pottery drawings, while remaining the format of choice for Web publishing.

The three main issues we could identify testing an archive of several hundreds of drawings are:

- a lack of native support for units (i.e. expressing the vector data in real units and not in pixels) can be worked around only with complex procedures, based on the <viewPort> element. To date, there is no straightforward way to deal with this issue, even though it could be addressed developing one or more plug-ins for the most used software programs;

- all programs have specific ways of dealing with image layers, that are not part of the native SVG 1.1 specification, but generally this is done through the <g> (group) element. This is not a severe limitation, in that single objects can be assigned an id attribute, that is unique in each file. This possibility is discussed below in the advanced workflows section;

- except for basic elements such as <title>, metadata can be expressed in different ways. While Inkscape uses Dublin Core and Creative Commons vocabularies, Adobe Illustrator follows the XMP standard. Both ways are acceptable and “standard” in their own way, but this discrepancy introduces another layer of complexity, making it sub-optimal to use different programs onto the same archive.

Finally, even though it is not strictly a limitation, SVG is a file format, so the expected structure of an archive is one file per drawing, as with other image formats. Collating more than one drawing in the same file is discouraged, even though this is common when preparing illustrations for print. In theory, nothing prevents saving serialised SVG data in a database (Jianting, Yan and Youguang 2004), to retain the relational structure often used in archaeological information systems.

6. Workflows

Based on the above outline and on first-hand experience with a growing collection of line drawings, several workflows have been devised, from simple to complex, with the aim of providing a step-wise approach to using SVG, rather than assuming a single given way to achieve the result of publishing drawings on the Web. The first and main point is that a liberal approach focused on the format is likely to find more acceptance than a rigid set of instructions.

6.1 Basic Workflows

Almost all pottery drawings are not born digital. As a consequence, digitisation through tracing is the rule, not the exception. The most basic workflow is to use SVG from the start to save digitised drawings, rather than other formats available in the vector graphic software used.

However, given how elementary this recommendation is, other steps can be added to the basic set. The following steps go in the direction of facilitating reuse, ensuring completeness of information and self-contained files:

- add a metric scale, with an explicit indication of the scale size;
• add metadata (author, title, description) using the “Properties” or “File properties” dialog of the drawing program; description and title can be as verbose as needed;

• again in the metadata editor, mark the file as distributed under a Creative Commons Attribution license ensuring ease of circulation to single files even when detached from the entirety of the archive.

Digitisation should be done according to the required graphical standard (e.g. section on left or right side, hatching colours, thickness of lines). An example is shown in Figure 1, comparing the hand-made drawing with the SVG counterpart. Depending on the complexity of the drawing, this procedure is not time-consuming and can be performed in five minutes. This workflow is also recommended in case of a conversion process from legacy or proprietary formats to SVG.

6.2 Advanced Workflows

The basic workflow guarantees a solid base, but keeps possibilities for further processing confined to metadata. Advanced workflows make processing of geometric data attainable, but they require more digitisation time.

The first kind of advanced workflow is inspired by programs that calculate vessel capacity (a recent overview in Engels, Bavay and Tsingarida 2009), and is based on a logical separation of the elements of the drawing: rotation axis, external profile, section are the three main elements to be considered, together with the metric scale. Drawing programs have a feature to give single objects an id attribute. The proposed convention is to adopt simple, unique and explicit labels like axis, profile, section, scale. An example in shown in Figure 2, with different colours highlighting the separate elements. With such data available inside drawings, it is possible to automatically extract size and measurements in metric units. Comparison of vessel profiles (e.g. as described in Modrzewska, Taroni, Pianetti 2010) is also technically possible and certainly not novel (Shennan and Wilcock 1975).

As explained above, handling of metric units is a weakness of SVG. The <viewPort> element can be used to work around this
problem, but to the best of our knowledge no software program is capable of providing an interactive management of this SVG feature. Thus, the only way to pursue this task is to manually edit the source of the drawing. For anyone who has ever written HTML, this is going to be a normal operation. XML-conforming editors should be preferred. The W3C Recommendation (Scalable Vector Graphics 1.1) contains detailed instructions on how to approach this task, but our experience showed that more than ten minutes may be required for each drawing.

The third and final step of an advanced workflow is the actual publishing of SVG drawings on the Web. While applications of any complexity can be developed for the task, it is certainly recommended to start with simple steps, taking advantage of the fact that modern web browsers can natively display SVG drawings. An HTML index with links to the SVG files makes it possible to publish drawings even on a minimal web infrastructure, as may be the case with low-cost shared hosting, only by uploading the containing directory via FTP (see also the catalogue application described above in §3). A basic example is shown in Figure 3.

### 6.3 Experimental Workflows

The importance of metadata was stressed in the above paragraphs. As a W3C standard, SVG can incorporate RDFa (Herman 2008), and is therefore capable of including semantic links (Berners-Lee 1998; see Boast and Biehl 2011, p. 150 for a critic of this approach) to other resources. As seen above for the `<viewPort>` element, this task is not possible without manual editing of the SVG source. Technically, this workflow is easily implemented even in the basic workflow. The experimental nature is more depending on the lack of available web resources to link to, for other drawings, excavation contexts of provenance, museum storage facilities and photographs (that, incidentally, can also be included in SVG files). Bibliographic references are at the moment the most promising example in this field, having undergone a significant exposure in the past few years, through the development of dedicated ontologies, in particular the Bibliographic Ontology (Glasson, D’Arcus 2009).

### 7. Conclusions

SVG is the recommended format for Web publishing and dissemination of pottery drawings, notwithstanding the variability of results given by different software in terms of XML content.

However, rather than being merely a technical matter, its adoption involves a shift in publishing habits, making place for dynamic knowledge items rather than static “snapshots” (Zanini, Costa 2009). While digitisation of existing archives is a daunting task, we hope that new work being carried out in the study of pottery will make a step in the direction outlined above, enabling further dissemination of high-quality digital archives.

This review of SVG for archaeological pottery is the result of an ongoing documentation project, available online at http://svg-pottery.rtfd.org/.

### References


Through an Urban Archaeological Data Model 
Handling Data Imperfection

Asma Zoghlami
University of Paris 8, France

Cyril de Runz
University of Reims Champagne-Ardenne, France

Dominique Pargny and Eric Desjardin
University of Reims Champagne-Ardenne, France

Herman Akdag
University of Paris 8, France

Abstract:
Similar to spatio-temporal information, archaeological information systems can be modeled using spatio-temporal modelling language such as PVL. Since archaeological data are largely imprecise, such imperfections must be taken into consideration. In order to deal with imperfections, a fuzzy set approach is used from modelling to querying. This article introduces a new archaeological data model in PVL and extends it in Imperfect PVL. Based on the modelling, a method for storing fuzzy data into a multivalent form will be proposed, and then the link between the classic (non-imperfect) data and the multivalent data will be exposed. Finally, the approach will be illustrated using some requests.

Keywords:
Excavation Data Modelling, Imperfection and Uncertainty Management, Database, Geographical Information System, Fuzzy Sets

1. Introduction

Since the beginning of the 1980s, there has been a need for excavation data storage, which has led to the deployment of information systems devoted to archaeological data. The secondary objective of using computational applications involved the creation of excavation site maps. However, in recent years, with the advent of new technologies and methods, it has become normal to use Geographical Information Systems (GIS) since they are better for both spatial data storage and analysis (Conolly and Lake 2006).

Today, projects handling excavation data and GIS tools are provided and utilized all over the world from an institutional point of view (Jordanian Web- GIS, Palestinian GIS) to a research point of view (Harrower 2010; Rodier et al. 2010).

City excavation data should be considered according to the triplet time-space-function (Rodier and Saligny 2007). In this triplet, there are seven spatial granularities from the stratigraphic units to the interurban area. Time is an integrable component of excavation objects, and it is generally a period of time. The function is, as in classic GIS, a part of the semantic information. The descriptive component is also formed by the information about materials, types of structure, etc. Therefore, as with spatio-temporal information, an archaeological information system could be modeled using spatio-temporal modelling language such as PVL (Perceptrory Software, Bédard et al. 2004).

Nevertheless, by regarding to the past, archaeological information remains in essence imperfect, and its quality should be taken into account from the information system modelling.
stage to the analysis stage. In this way, data imperfections may be identified, characterized, memorized and thus rendered queriable by an archaeological GIS.

Thus, an adaptation of classic GIS modelling language is used in this article in order to model archaeological data and the imperfections inherent in that data. Based on this modelling, a new database approach will be introduced, and some simple spatial and temporal queries will be shown.

This article is structured as follows: Section 2 shows the excavation information system modelling using PVL. Section 3 shows an adaptation of the model and its database consequences in order to represent data imprecision. Section 4 introduces some possible queries. The final section (5) is devoted to the conclusion.

2. Archaeological Data Model

The handling of urban excavation data is the primary issue in understanding the past and delivering this knowledge to today’s citizens. For this reason, it is essential to store and be able to visualize archaeological data.

In urban excavations, there are two principal studied spatial objects. On one hand, the studied site is essential because it is the location of the phenomenon, and may provide information about the global past. On the other hand, there are archaeological entities that are present in an excavation site. Those entities correspond to the excavated objects that represent parts of the past structure, such as walls, rooms or buildings.

An urban archaeological entity is structured over seven spatial granularities from the stratigraphic unit to the interurban area. On a more global scale, an entity may be composed of objects and entities representing data on a more local scale. The granularity of objects is obtained by defining classes that specializes the archaeological entity class.

As only excavation data are considered, an archaeological entity has a spatial shape, a function (wall, oven, house, etc.), and a period of activity – the period when the object was used for its previous function. The kinds of archaeological entities (different granularities) handled by the developed information system are (according to Rodier et al. 2010):

- Stratigraphic units: these are, in general, an occupation layer within a building. This can be a wall, ditch, etc. It may contain artifacts, which are objects formed by humans that present an archaeological interest (e.g. coins, pottery).
- Facts: these are voluntary structured sets, and can be separately isolated and studied. One or more stratigraphic units can be attached to the fact.
- Structures: these are a collection of individual archaeological facts, which constitute a coherent set that can be interpreted as a complex fact, e.g. a ceramic oven.
- Building blocks: a coherent set of entities at a higher level (e.g. a house).

The archaeological data model is developed using the language PVL (Bédard et al. 2004) primarily based on UML extended with spatial and temporal stereotypes. The model, presented in figure 1, corresponds to the concept of an archaeological site and an archaeological entity.

PVL spatial pictograms allow us to define the geometry chosen for each one of them. Archaeological entities and sites are considered to be areas while the geometric type representation of the artifacts and documentation represent the point.

In order to represent dating that corresponds to possible periods or centuries,
temporal entities, which have a sustainable existence like the sites, archaeological entities or the artifacts, are represented by a pictogram indicating a time interval. The documentation class, which has an instantaneous existence, is represented by a pictogram indicating its creation date.

In addition to these entities, we will also consider the spatio-temporal documentation class. The documentation class, in addition to having a location and date, allows the reference to be stored on entities.

The archaeological entities, documentation and artifacts have also been located. Thus, they are linked to the georeferencing class that saves their geographical coordinates. However, the sites are localized through their postal address. Figure 2 highlights the spatial structure of our excavation information system.

These classes, with the exception of the site class, have descriptive characteristics (such as dimension, height, width, thickness, length, etc.) in addition to spatial and temporal characteristics. The archaeological entity is linked to the material class that provides a description of all the materials composing it.

The artifacts also have their own description in terms of materials.

The notice class allows the context of the concerned objects to be described and characterized by associating original documents or parts of their contents to them. Keywords give the archaeologists the choice of describing objects using either a structured language and a free language. Information about artifacts and the preservation of documentation (storage addresses, storage conditions, etc.) is stored in the storage class. The sample class includes data concerning samples of archaeological materials such as ceramics, wooden objects, etc. Figure 3 shows the global structure of the system.

Several archaeological data models have been presented, such as those introduced by (Pfoser et al. 2007) or (Rodier et al. 2010). The former is well-documented for data storage during excavation, while the latter is used for data analysis. Both, however, have yet to be implemented as multi-site storage systems that
store final excavation data, which is one of the goals of our system.

The proposed archaeological model handles common scenarios where an archaeological object has had multiple, different functions over time, using the aggregation relation between entities. The object is then modeled as a complex entity that is composed of more simple objects having a definition in accordance with one object – one function – one location – one period of activity. Moreover, the spatial relations (overlaps, intersects, includes, etc.) may assist in identifying the reuse of objects.

This approach is strongly inspired by the system currently used for storing excavation data about Reims known as GISSAR (Pargny 2008). This system allows for the handling of archaeological data but does not allow us to consider data imprecision (imperfection). As archaeological data are often uncertain or imprecise, this aspect must be considered. That is the focus of the next section.

3. Handling Imperfect Data: Fuzzy Conceptual Data Model

3.1 Imperfection in Excavation Data

We discovered that imperfections in the archaeological data handling process start to occur during the acquisition of data. This continues through the representation and management of data, only ending during data visualization. Indeed, the time interval between the excavation and the activity period of an object implies some imprecision/uncertainty: partial destruction, soil movement, an estimation of the activity period, an estimation of descriptive attributes, etc. Several types of excavation imperfections have been introduced in (de Runz et al. 2011).

Quantitative information is often defined using qualifiers. For instance, the type of stone shape in a wall tends to use qualifiers such as long, thick, thin, and large or even moderately thick, very thin and a little bit large. This is meaningful in the case of artifact storage. This semantic imperfection is also related to the composition of an archaeological instance. For example, one may wonder whether an instance contains iron or marble, or neither. It is thus related to the archaeological instance function which can end up being uncertain.

Time information is related to the activity period of the object. We often have imprecise information concerning the existence of a historical event or an object because its description takes the form of sentences such as “It happened at the beginning of the Second Century”, “It was during the middle of a period” or “It happened during the Third Century”. Sometimes, this description may also take the form of expressions such as “It was at the very beginning of a century” or “It was right in the middle of a period” and so on.

Spatial data includes information about the geometry and the location of an instance on Earth. In classical approaches, spatial instances are usually represented as objects with well-defined boundaries (points, lines, regions) even if, by their nature, they are vague, such as the boundaries of a stratigraphic unit or an archaeological site. The location of an instance can be made by using reference points, benchmarks, using expert positioning, and through old repositioned maps. Thus, the georeferencing of an instance is usually related to a lack of precision.

3.2 Handling Imperfection through Information System Modeling

Dieter et al. (2007) proposed a system prototype dealing with the spatial uncertainty of finds and excavation diaries, which records an excavation’s progress. This spatial uncertainty is considered in the data model, through the visualization of uncertain positions and the query results. The position is implemented
either as a point or as a polygon, depending on the object’s type.

The previous archaeological model, for instance, does not allow for the handling of imprecision. However, according to (Shu et al. 2003), the uncertainty of a geographical entity can be modeled through consideration of the uncertainty in its spatial, temporal and thematic attributes. This tool focuses on uncertain spatio-temporal data in the conceptual data model MADS, but it does not provide a suitable representation for more general cases of imperfection.

Other approaches have proposed visual pictograms for the management of imperfection, such as the fuzzy data model of (Ma et al. 2010). There is also a UML extended model handling fuzzy data. In this model, different fuzzy levels are introduced into the class diagram (fuzzy class, fuzzy association, fuzzy aggregation, fuzzy generalization, etc.), but the spatio-temporal dimension is not considered. In the spatio-temporal context, pictograms handling imperfect spatio-temporal data were introduced in (Miralles 2006) and (Zoghlami et al. 2011).

In this article, we propose to model and implement imperfect archaeological data using the last approach. This approach allows us to represent imperfection by adapting PVL pictograms: imprecise spatial and temporal data are both represented by traditional symbols with dashed rectangular outlines, while attributes that are imprecisely defined are modeled by introducing the keyword IMP before the attribute’s name.

In the PVL archaeological data model, we distinguish different levels of imperfection. The first level is concerned with imprecision of descriptive characteristics (dimension, composition, etc.).

In fact, we describe the dimension with fuzzy predicates such as thick, high, long, etc. Moreover, we may have uncertainty about the materials that constitute the entities and samples, and uncertainty about the superposition relationships between the stratigraphic units.

In regards to the documentation, on the one hand, reliability presents a problem. The document’s reliability, and the amount of confidence we put into the document should be considered in terms of the document’s originality, content, or the relevance of the author.

The second level of imperfection is related to time. Temporal features of archaeological entities correspond to time periods where the considered objects were active. This dating has a lack of precision, since we cannot precisely identify the two terminals of the time interval.

The final level is related to space, namely the geometric shape of space objects that may have fuzzy boundaries in addition to the imprecision of their georeferencing.

To handle the first level of imperfection the keyword IMP is introduced and placed in the dimension class before imperfect attributes such as length, width, height and thickness. This article will not consider semantic imperfection; it will only handle the attribute fuzzy quantitative values. To handle the second level of imperfection, we use PERCEPTORY classical temporal pictograms with dashed outlines to express the time imprecision. Objects having an instant existence are represented by a dashed pictogram indicating an imprecise date. The main issue with objects that have a sustainable existence, like archaeological entities, relates to their temporal boundaries. These objects are represented by a dashed pictogram indicating a time interval imprecision.

For spatial imperfection, the fuzzy boundaries of the archaeological sites and archaeological entities are modeled through a polygonal spatial pictogram with dashed
outlines. The artifacts and documentation are modeled through point geometry with dashed outlines. The georeferencing imprecision is also modeled through a point pictogram with dashed outlines. Figure 4 presents an extract of the Imperfect PVL Diagram highlighting the three levels of imperfection in an archaeological context.

As the fuzzy set theory introduced by Zadeh (1965) is a good approach for dealing with imprecision and uncertainty, we choose to model data according to fuzzy sets in this article. The database is thus adapted to store and query fuzzy spatio-temporal data, as explained in the next section.

3.3 Fuzzy Spatio-Temporal Database Structure

As the main goal involves the storage of fuzzy spatio-temporal data, we chose to build a system that organizes the information into several layers. We implemented a first layer, called the data layer, that contains geometric data such as the shapes and locations of archaeological instances, descriptive data referring to all the object descriptive attributes and temporal data related to all the temporal classes mentioned above.

A second layer, the multivalent layer, is linked with the data layer. This allows for the handling of imprecision and uncertainty through multivalent representation.

The principle of the multivalent approach lies in the introduction of several truth values that modulate information in order to focus on the imperfection of natural language. According to some studies (Revault d’Allones et al. 2007), seven is the maximum number of truth values required to use knowledge well. Thus, linguistic expressions, such as “very little”, “a lot”, “medium”, etc., can be used (Akdag et al. 2008). In order to illustrate our approach, six truth values will be used. However, our approach allows fuzzy sets to be defined using two to seven α-cuts. The imperfection layer is followed by a meta-layer representing the metadata. This layer usually contains data sources, data identification, data quality, spatial representations, spatial references, specific ontologies, database schema and any other useful characteristics that may qualify the data. At the implementation level, we ensure communication between the three layers.

We are interested in the imperfection layer that handles semantic, spatial and temporal uncertainty through a multivalent approach.

In regards to a quantitative qualifier for imperfection due to the use of natural language, all the fuzzy attributes are connected to an imperfect table that stores the imprecise information as a multivalent set of values via an intermediate table which references their fuzzy set id. Figure 5 illustrates an example of managing imperfect attribute thickness (attribute 1) and length (attribute 2). The fuzzy labels “thick” and “thin” correspond to the first
attribute thickness while the fuzzy labels “long” and “short” correspond to the second attribute length.

For the spatial data, we create a space imperfection table which includes all the geometric shapes corresponding to archaeological sites and entities as fuzzy polygons. In addition, all the geometric shapes corresponding to the artifacts and documentations are included as fuzzy points. All spatial queries will refer to this table. Figure 6 represents all the fuzzy forms corresponding to the archaeological site “shape_id 233.”

For temporal data, all the temporal entities are connected to a time imperfection table that stores all the semantic expressions referring to the date as a multivalent set of values. Connection with this table is possible through an intermediate table indicating the fuzzy set identifier for each dating reference. For example, for reference dating that indicates the middle of the 2nd Century, we can have different truth values that correspond to the time interval chosen (Fig. 7).

In this approach, the indexation of data is given according to the index of the objects/attributes modeled as imperfect. The relation between the object/attribute and its imperfect forms should therefore be viewed as a link-id relation.

Using the previous principles, we are now able to build an operational GIS devoted to archaeological information. A case study with some query examples is given in the next section.

4. Illustration

The following examples of simple queries are shown on simulated data. Our excavation data spatio-temporal configuration is similar to this situation.

4.1 Imperfect Spatial Query

Here, we consider a request aiming to locate shapes that roughly correspond to the geometry of site JO 88. The retained shapes are those that have a membership degree to the specified site that is greater than or equal to the value of 0.4. Visualization of the previous results in a Quantum GIS shows a set of shapes surrounding the site that have various membership degrees. These membership degrees are between 0.4 and 1, which are values that correspond to archaeological site JO 88 (Fig. 8).

This simple query allows the selection and visualization of a multivalent shape higher than a degree. This is important for the presentation of uncertain results.

4.2 Imperfect Temporal Query: Anteriority

Here, we consider a request aiming to locate shapes that roughly correspond to the geometry of site JO 88. The retained shapes are those that have a membership degree to the specified site that is greater than or equal to the value of 0.4. Visualization of the previous results in a Quantum GIS shows a set of shapes surrounding the site that have various membership degrees. These membership degrees are between 0.4 and 1, which are values that correspond to archaeological site JO 88 (Fig. 8).

This simple query allows the selection and visualization of a multivalent shape higher than a degree. This is important for the presentation of uncertain results.
or absolute definition for the anteriority (and posteriority) between fuzzy periods or dates. Our idea is to moderate the decision (anterior or not) by providing it with a confidence index. For that, we use an index introduced in (de Runz et al. 2010) that evaluates the anteriority between two fuzzy periods. Figure 9 shows a sample of the output. Some other fuzzy temporal indices may also be used (Dubois et al. 2003; Schockaert et al. 2008).

4.3 Imperfect Spatio-Temporal Query

We consider a request aiming to find entities that satisfy the following conditions:

- Their activity period is the 2nd Century (with a degree of at least 0.4).
- Their shape belongs to the site “PC 87”.
- The final degree must be at least equal to 0.8.

This request corresponds in fuzzy logic to:

\[
\text{ActivityPeriod}(x) \sim 2^{\text{nd}} \text{ Century} \geq 0.8 \ \text{AND} \ \text{Shape}(x) \sim \text{PC 87} \geq 0.8.
\]

The visualization of a query combining spatial and the temporal imperfection is illustrated in figure 10. This figure illustrates an example of a query returning entities that have an activity period in the 2nd Century and that belong to the site PC 87. According to this example, an entity having an identifier of 356 is the only one that satisfies the two conditions.

We presented three cases where a GIS built to handle archaeological data may be used, considered their imperfection (in our examples, their imprecision).

5. Conclusions

In this paper, we developed an approach that allows for the handling of imperfect data in an archaeological context. The first step was the introduction of a model for building an archaeological GIS. In the second step, the imperfection – and, in particular, the imprecision – was identified in the model in order to build an information system able to handle the imperfection of archaeological information. The third step consisted of building a fuzzy spatio-temporal database (with a multivalent representation of fuzzy data). The last step indicated queries for extracting data and their imprecision in temporal, spatial and spatio-temporal contexts.
Even though several approaches/theories exist for dealing with imperfect data, we focused on fuzzy approaches because they are useful for imprecision modelling and allow for the simple representation of natural languages. This is a requirement for our system. In fact, the fuzzy set theory was well used in our context and demonstrated its potential through both archaeological data analysis and mining processes.

Nevertheless, our system permits the storage of other kinds of data representation (probability sets, multi-sets, rough sets, evidential data, etc.). In this case, the queries should be adapted in accordance to the chosen theory, and may be an aspect of possible future improvements that we would study.

Our future work will be devoted to the utilization of Allen's relations (Allen 1983) and their fuzzy definitions in both time (Dubois et al. 2003) and space (Salamat and Zahzah 2010).

References


Rodier, X., and L. Saligny. 2007. “Modélisation des objets urbains pour l’étude des dynamiques urbaines dans la longue durée.” In SAGEO’07 Spatial Analysis and


Guerrilla Foursquare: a Digital Archaeological Appropriation of Commercial Location-Based Social Networking

Andrew Dufton  
Brown University, USA

Stuart Eve  
University College London, L - P : Archaeology, UK

Abstract:  
One aspect of the emerging field of digital archaeology involves the use of geo-technologies to create and disseminate location-based archaeological information. Although archaeological projects often lack the resources necessary for tailor-made applications, existing services can also be repurposed for archaeology at relatively low cost. A specific case-study using the foursquare application in central London highlights the potential for (mis)use of existing services. Users of foursquare ‘check-in’ at various locations on mobile devices to access recommendations, locate friends, or gain digital control of venues. Through the inundation of the foursquare service with archaeological sites within the Roman city, the project created a palimpsest of past urban landscapes reaching a broad, non-archaeological audience. Public users can explore this landscape as well as contribute additional layers of narrative, or ‘tips’, ultimately creating a digital application evolving beyond the scope of the initial project.

Keywords:  
Geo-Location, Social Media, Mobile Applications, Roman, London

1. Introduction  
The rapid and ongoing development of increasingly complex mobile technologies has had a profound impact on the way connected societies interact with built and natural spaces. When walking through a city or a rural space it is now all too easy to reach into one’s pocket, pull out a smartphone and have access to all of the world’s knowledge at the touch of a button. When walking past an archaeological site, we can immediately look up the history of the site on Wikipedia and even in some cases access the original site excavation reports. Due to its GPS chip and cell-tower triangulation, the smartphone knows where we are and can deliver context-aware information directly to us in a timely fashion (Brown 1998; Coppola et al. 2005).

Aside from an occasional tendency toward overzealous adoption of the latest “archaeological toys” (Zubrow 2006), the integration of these new technologies into a digital archaeological practice can not only add a sense of legitimacy to academic projects but also capture the imagination of the wider public. In particular, the geo-location services inherent to most mobile applications can become a valuable tool in the creation and dissemination of location-based archaeological information, to both academic and non-academic audiences. Despite this potential many archaeological projects lack the resources or expertise necessary to create effective, tailor-made mobile applications. As a result, mobile archaeological applications have become limited to only the ‘have’ projects or museums, with the ‘have-not’ projects unaware of or discouraged from adopting the latest digital technologies.

Yet in many situations existing commercial services fulfilling a similar purpose could, at relatively low cost or effort, be successfully repurposed for archaeological uses. As a case-study in the use/misuse of commercial...
services, the foursquare location-based social networking application was appropriated to add archaeological information on the remains of the Roman city at a variety of sites across central London. Through tracking the interactions of a broad, non-archaeological audience with a digital reminder of buried archaeological landscapes, some ideas on the wider potential for reuse of existing social networking in the field of digital archaeology will also be discussed.

2. Mobile Applications and Archaeology

Archaeology and cultural heritage lend themselves well to a mobile approach. We walk, live and work within a palimpsest of cultural heritage and people are interested in their past and the past that surrounds their daily lives. This paper will not discuss the advantages and disadvantages of the use of handheld devices in a museum or archaeological context per se, but rather will concentrate on the form of delivery of information through the devices themselves. For a fuller discussion of the various issues involved with using handheld devices see Tallon and Walker 2008. A number of mobile applications have developed that attempt to share knowledge with the casual smartphone user. These applications (or apps) can be divided into two different types: apps or platforms that allow users to upload archaeological information and apps specifically designed solely to deliver archaeological material.

Platforms that allow direct uploading of archaeological information include Layar (http://www.layar.com/), HistoryPin (http://www.historypin.com/) and Junaio (http://itunes.apple.com/gb/app/junaio-augmented-reality-browser/id337415615?mt=8). Each of these platforms allows users to create ‘channels’ or ‘layers’ of historical or archaeological information that can then be uploaded and made available to other users that choose to access those channels. For example, at the time of writing (June 2012) Layar had 5 archaeological channels available, 3 of which were virtual tours of archaeological sites. The user normally downloads these types of channels when they arrive at an archaeological site; the mobile application guides them around the site and supplies context-aware information. This contributes to the user’s knowledge of the site and could also be a cost-effective way to replace traditional site guide equipment (bulky audio tours, etc.). Downloading is an active decision (usually prompted by signs at the site entrance) and therefore the channel is usually only downloaded by users who have already shown an interest in the archaeology.

Examples of specifically designed apps include the Museum of London’s ‘Streetmuseum’ (www.museumoflondon.org.uk/Resources/app/you-are-here-app/home.html), Melbourne museum’s ‘Please Touch the Exhibit’ (http://museumvictoria.com.au/melbournemuseum/discoverycentre/pleasetouch-the-exhibit/) and the Phillips Collection app (http://itunes.apple.com/us/app/the-phillips-collection/id402056956?mt=8&affId=1503186). These apps have a custom design with the aim of adding value to the collections of the museums, while also encouraging people to visit the institutions themselves. The downloading of the app is again a deliberate act and a demonstration of a prior interest in the museum collection and heritage itself. The Streetmuseum app uses the GPS and video feed of the smartphone to ‘overlay’ the collections of the Museum of London onto London streets using the Augmented Reality view. Essentially this enables the user to walk around London and ‘see’ old photos and prints in the location that they depict - and compare them to the modern view. The app also supplies additional place-specific information from the Museum’s archives. According to Jeater the deployment of the app has been successful in raising the profile of the Museum, particularly amongst young adults (Jeater 2012, 73).

The above examples of specifically archaeological applications making use of
mobile technologies demonstrate some of the advantages of such an approach to public outreach and engagement. However there are also a number of drawbacks to bespoke solutions. In the case of the Museum of London’s Streetmuseum £48,000 was spent on development (Jeater 2012, 73). The media firm working with the museum, Brothers and Sisters Creative Limited, also matched all museum funds with additional free development time. In return, the firm has maintained the intellectual property of the Streetmuseum technology, which could theoretically be marketed to other heritage clients. Although this partnership has been effective in the high-profile example of the Museum of London, it is unlikely most archaeological projects or museums would have the resources and institutional clout needed for similar success.

Yet beyond the cost and licensing implications of this development paradigm there are further, perhaps unintended, consequences to archaeology or heritage specific applications. Platforms created specifically for archaeological purposes rely on a previous interest on the part of users in both discovering and downloading any new application. A targeted campaign to publicise a new application can help to mitigate this limitation; in the example of Streetmuseum, positive press coverage helped the application reach over 40,000 iOS downloads within the first two weeks of release (Jeater 2012, 76). Nonetheless, by downloading the Streetmuseum app these new users demonstrated a direct interest in the museum and its archives. The outreach potential of Streetmuseum, or any similar mobile applications, does not include the broader public who may or may not otherwise demonstrate any interest in interacting with archaeological or heritage information.

3. Guerrilla Foursquare

A case-study using the foursquare location-based social networking service demonstrates the potential for repurposing existing commercial services for archaeological applications. Foursquare was launched as a social media platform in 2009, first available to iPhone users, through SMS messaging, or using online browser access (Van Grove 2009). According to their website, the general purpose of foursquare is to “Keep up with friends. Discover what’s nearby. Save money. Unlock rewards” (Anon. 2012a). Using a loosely-formed game concept, users score points for ‘checking-in’ at known locations and compete against other members of their foursquare network. Additional points may be awarded for checking-in at new locations, becoming the first within a friend network to discover a new place, repeated daily or weekly check-ins at the same foursquare entry, or distance travelled between stops. The user with the most check-ins at a given location in a two month period is named its ‘mayor’, with many commercial organisations or public places encouraging frequent visits with discounts or rewards for the reigning mayor. In addition, a series of general badges are awarded to users for specific accomplishments, such as the ‘I’m on a boat’ badge for a check-in at a nautical location or the ‘Swarm’ badge for checking-in with over 250 other foursquare users.

Initial adoption of the foursquare network was limited to a small population of trendy smartphone users in larger urban centres (Van Grove 2009), likely due to the greater number of locations in cities and, with a larger user base, more opportunity for social interaction. Foursquare’s creators also actively cultivated partnerships with commercial providers such as American Express and Dominoes Pizza (Kimberley 2010; Overby 2011), as well as public institutions. As with many other social media platforms, the user base of foursquare increased greatly after the service began offering integration with the popular social networking platform Facebook, and according to the latest figures there are over 20 million worldwide foursquare users (Kessler 2012).
For this case study, foursquare records in central London were ‘archaeology bombed’ using recycled data from the Greater London Historic Environment Record (GLHER). From the over 9600 Roman period entries, a total of 48 GLHER sites from Roman London were specifically selected accounting for both excavated and standing archaeological remains and including evidence of residential, commercial, civic, military and ritual activity. To comply with foursquare’s existing functionality, the detailed GLHER entries were distilled into brief, accessible summaries of 200 characters or less summarising the Roman presence discovered during modern construction works at each site. These summaries were then added as tips at each of the selected sites, making them visible to foursquare users checking-in at the site or within the general vicinity (Fig. 1). The selected Roman locations also formed a public foursquare ‘list’, Roman London (underground), collating the sites for easier tracking of user interaction and facilitating cross-discovery of the selected sites by individual users (Fig. 2).

In some cases, existing foursquare locations at the site of archaeological remains were given additional archaeological information; for other examples a new foursquare site was added specifically for buried archaeological remains. Each site was then assigned a type from an extensive list of site types provided by foursquare (Fig. 3). For the purposes of this discussion these specific site types have been grouped according to the following broad categories: historic sites, commercial premises, office buildings, public space, museums, public transport, and other.

4. Tracking the Users

During a three month trial period there were over 2121 total check-ins at the 48 selected sites, for an average of approximately 35 check-ins per day. However, looking at the breakdown of check-ins by site category shows a high degree of variation for users visiting different site types (Fig. 4). For example, commercial and office buildings had almost 750 check-ins each during the three month period yet, in stark contrast, the 16 places categorised as ‘historic sites’ were visited a pathetic total of twice over this time. Public spaces, museums, and transport hubs fell somewhere between these two extremes, with around 200 check-ins for each category.
Aside from the obvious disinterest in historic sites, looking at the total number of sites in each category provides some further insight into user patterns. For example, the presence of 12 commercial sites in the selected sample, the highest category after historic sites, makes the resulting higher number of check-ins less surprising. Yet the single museum on the list, the Museum of London, achieved 229 check-ins to rank as the most popular single destination. To better visualise the relationship between number of sites and total check-ins for each category, a ratio of the percentage of sites to the percentage of check-ins was created (Fig. 5). By this metric, those sites with a ratio greater than one are attracting a proportionately higher number of users than those with a lower ratio. Perhaps unsurprisingly, the highly trafficked Museum of London, as well as the transport hub at Bank Underground Station, were visited by a significantly greater number of users than more mundane office buildings or smaller public spaces. At a ratio of 0.0028, the failure of historic sites to engage foursquare users becomes even more marked. The number of users acknowledging they have ‘done’ the archaeological tips was also recorded at the end of the trial period. Tips were only marked as completed by four distinct users at two of the 48 foursquare locations. An additional three users subscribed unsolicited to the Roman London (underground) foursquare list.

A number of general observations can be made from the above statistics. First, there is a marked disinterest among foursquare users for the historic sites created specifically for buried or excavated archaeological material. The almost total absence of check-ins at these sites could be representative of a wider disinterest on behalf of the public in remains discovered during commercial works divorced from the initial excitement of discovery. More likely, however, this lack of activity is caused by broader patterns in foursquare use. Historic sites containing buried archaeological material, or those sites previously excavated for commercial purposes, are often invisible at street level. Foursquare users primarily check-in at the places they are already visiting throughout the day, rather than using the service to explore unknown locations. As a result, most users in central London likely never discovered the new archaeological listings, let alone expressed a direct presence or interest by checking-in or completing a tip, respectively. In order to be effective, information relating to archaeology or heritage must be associated with sites people are already visiting or aware of, rather than attached to less tangible underground locations. Directly related to these observations, those sites with the most regular public traffic are also those recording the greatest number of foursquare user check-ins, and so hold the greatest potential for discovery-based learning.
by a wider audience. For example, a Starbucks coffee shop at 74 Cornhill was one of the most popular commercial check-in locations. By adding details of the Roman forum foundations excavated during construction, city workers stopping for a morning coffee and checking-in at Starbucks are also stumbling across unexpected details of the buried past.

Moving from check-ins to tips leads to some other interesting observations of foursquare user behaviour. At first glance, the minimal number of users acknowledging the GLHER summaries seems to suggest very little actual interaction with the archaeological material. Yet looking at the adoption of other tips at these sites suggests a more systemic problem with using tip completion as a means of tracking public engagement. Even established organisations received very little uptake of tips at many locations. For example, at the time of this study only seven users completed the popular publication TimeOut London’s tip at the Museum of London, outlining the information contained within the museum’s galleries. Only four foursquare users had acknowledged a similar tip by the World Service of the British Broadcasting Corporation left at Fleet Street. Compared to these well-known media outlets, the four tips completed for archaeological remains, created by an individual user, seem almost optimistic.

To understand the seeming inefficiency of tips, some consideration of the foursquare mobile interface is needed. Up to three tips appear for a given site, under the main listing information (Fig. 1). Therefore the three most popular tips for any location are visible to all users checking-in. However, to ‘do’ a given tip requires tapping the screen first to read the full tip listing, then a further two screens to acknowledge completion and confirm. This unnecessary three-step process discourages most users from actually recording any tips they may have completed, making it difficult to accurately track actual user interaction with tips clearly visible at foursquare sites.

Looking to the most popular tips for guidance, some general guidelines for attracting users can be established. For example, the History Channel recorded the most popular tip at many of the London sites, including almost 200 users completing a tip regarding the historical foundation of the Museum of London. The tip formed part of a wider list formed by the History Channel, with over 300,000 followers, of historical details of London sites. Subscribers to the list receive pop-up notices from the History Channel whenever they check-in near a historical location, resulting in both a greater degree of exploration by foursquare users and a more concrete means of tracking public interaction. From the History Channel example it is clear that the added legitimacy of institutional or commercial affiliation is best exploited not by tips alone, but through the construction of a community of users looking to add value to their foursquare experience.

5. Foursquare the Future?

During this short paper we have attempted to show the possibilities of appropriation of a commercial mobile platform for archaeological means. By choosing a platform already in heavy use, we have the potential to reach a vast audience at a very low cost – in this example only the amount of time it takes to import the data from the HER. From the data gained during this trial period it would appear that foursquare users don’t actively seek out historic sites, and so adding archaeological sites as standalone features is not effective. Instead we employed guerrilla tactics and added tips and information to existing well-visited sites, as in the Starbucks example discussed above. Users who check-in to their normal coffee shop or office will accidentally discover that the building has an archaeological history, even if they weren’t actively seeking out this information. Obviously more work needs to be undertaken to follow the long-tail of this discovery; for example do the users then share this new information with others? Do they
download a specifically designed archaeological app to follow up on their chance discovery? Does it pique their interest enough to decide to visit the local museum at lunchtime? The upfront investment in the re-appropriation of services such as foursquare is so negligible – compared to that of a custom-made app, for instance – and the potential for new audiences for archaeological information so great that we believe these guerrilla tactics have a real place in the public archaeology movement and should be employed whenever possible.

As an appropriate post-script, a new version of the foursquare application was revealed to users in the early weeks of June, 2012. The change is described by foursquare as a complete overhaul designed to track user experiences and preferences and take better advantage of social discovery networks (Anon., 2012b). Archaeological applications could be well served by this shift in focus, allowing a greater degree of discovery-based learning within the existing foursquare game architecture. However this restructuring also serves as a sombre warning: using commercial applications often involves a degree of impermanence that could be problematic for some projects. Yet for most these dangers are far outweighed by the benefits, making ‘guerrilla’ appropriation of commercial mobile services an effective way of repurposing data and communicating archaeological material to a wider audience.

References


Conceputalising eScience for Archaeology with Digital Infrastructures and Socio-Technical Dynamics

Teija Oikarinen
University of Oulu, Finland

Helena Karasti
University of Oulu, Finland and Luleå University of Technology, Sweden

Abstract:
The paper introduces the concept of Digital Infrastructure (DI) and the associated notion of Socio-Technical Dynamics to archaeology. DI conceptualization of technology goes beyond the ‘technology as tool’ notion prevalent in archaeology by acknowledging the process of digitalization and socio-technical dynamics in integrating technologies into expanding configurations. We propose DIs for analysing archaeology in the eScience era, and as an illustration we discuss the paradox of control and change that concretizes for example in information models and standards. Standards have a central role in enhancing interoperability and ensuring the future use of data. The socio-technical dynamics of DIs recognizes standards together with technologies and users as networked constituents of DIs. In cultural heritage and archaeology, standards are emerging as global infrastructure components. We discuss how technology and standards as controlling mechanisms have implications for archaeological research practices and scientific procedures. Research into these mechanisms can guide designers, managers, and policymakers.

Keywords:
Digital Infrastructures, Socio-Technical Information Systems Research, eScience, Socio-Technical Dynamics, Digitalization

1. Introduction

eScience (a.k.a. cyberinfrastructure in US) denotes ‘the layer of information, expertise, standards, policies, tools, and services that are shared broadly across communities of inquiry but developed for specific scholarly purposes’ (American Council of Learned Societies 2006, 1). eScience promises to transform research through: (1) community-wide and cross-disciplinary collaboration, (2) computationally driven collections, and (3) representation, analysis and integration of digital data. eScience is a scientific research area that uses networked digital information technologies to create and enhance cooperation, data sharing, and dissemination of research outcomes between disciplines or domains. (Ribes and Lee 2010, 231–233). Thus, Digital Infrastructures form the basis and are one of the main research topics in eScience.

Large-scale Digital infrastructures (DIs) (Tilson, Lyytinen and Sørensen 2010b) have begun to evolve also in the field of cultural heritage management (International Council of Museums 2012) and archaeology (for example McManamon and Kintigh 2010; Watts 2010). ‘Digital infrastructures can be defined as the basic information technologies and organizational structures, along with the related services and facilities necessary for an enterprise or industry to function’ (Tilson et al. 2010b, 748). DIs and related socio-technical research in eScience can be seen to offer a new conceptual framework to analyse and theorize complex phenomena, processes, and factors, which are simultaneously social and technical, as well as global and local (Edwards, Jackson, Bowker and Knobel 2007, 5–6). This article argues that Digital Infrastructures and the associated notion of Socio-Technical Dynamics are in demand and can be applied also for archaeology.
Recent literature reveals that technology in archaeology is frequently understood as individual applications or information systems for research and analysis (Huggett 2012). Huggett suggests that there is a need to reflect on the broader implications of technology in archaeology, ‘to think beyond the tools, not just think about tools’ (Huggett 2012, 204–213, emphasis added). Huggett and Ross (Huggett and Ross 2004b, cited in Huggett 2012, 204) identify three levels from which to approach the discussion of archaeological computing: (1) ‘the specific applications [level], i.e. its implementation and use, (2) the origins and prospects of large systems, and (3) broader implications of information technologies within archaeology and how they are integrated to the subject’. According to Huggett and Ross, levels 1 and 2 are widely explored in archaeological publications, but level 3 is rarely discussed. Huggett suggests Science and Technology Studies (STS) as an inspirational and ‘provocative source for archaeologists’ who are trying to ‘go beyond the tools’, to find new research questions relating to the ‘very purpose of considering the development, effects, and implications of new technologies’ (Huggett 2012, 212).

While we align with Huggett’s suggestion and use socio-technical Information Systems Research (ISR) and STS inspired research for this article, we argue that merely ‘going beyond the tools’ as technology notion does not seem to be enough in the context of eScience. Rather eyes should be targeted on the large-scale infrastructures and networks formed by information technologies, on the data produced and mediated therein, and on the intertwined social and technological relations and complex contexts formed in these processes.

Another stream of research for studying complex socio-technical systems is data-driven computational modelling (Vespignani and Santorras 2007; Vespignani 2012) where dynamical processes in complex socio-technical systems are modelled, for example focusing on the structure and stability of the Internet (Vespignani and Santorras 2007). In this approach there is an assumption that the complexity of socio-technical systems can be quantified, their structure and dynamics delineated and mathematically described, a set of relevant factors identified, and data about them generated using the increasing availability of computer power and informatics tools. The aim is not only to understand the complexity and dynamics of the studied bounded phenomena but also to predict and control their dynamics (Vespignani and Santorras 2007; Vespignani 2012).

The body of STS and socio-technical ISR inspired research holds, in turn, that digital infrastructures as complex socio-technical phenomena cannot be defined through a distinct set of functions (unlike specific IT systems), or strict boundaries (unlike applications), in contrast, they are seen relational in nature and characterized by dynamism and longevity (Tilson et al. 2012, 1324–1325). There is an assumption that DIs can be defined as ‘shared, unbounded, heterogeneous, open, and evolving sociotechnical systems comprising an installed base of diverse information technology capabilities and their user, operations, and design communities’. (Hanseth and Lyytinen 2010, 4 cited in Tilson et al. 2010b, 748–749). DIs are understood through qualitative analyses of processes of embedding capabilities and standards in organizational practices, which enable new social behaviors and/or regulations (Edwards et al. 2007, 5–6). We will further define and discuss DIs throughout the article, while this brief introduction suffices to highlight that the two approaches to studying complex socio-technical phenomena are clearly very different in their assumptions, yet likely both important at this early state of limited understanding of large-scale DIs as complex socio-technical phenomena.

We will now proceed to elaborate the definition of DIs, discuss their past, present and potential future(s), and present the notion of
Socio-Technical Dynamics. We will also discuss how this relates to archaeology, and conclude by summarizing the contribution of the used viewpoint in this article for archaeology.

2. Defining Digital Infrastructures

Basically, Digital Infrastructures (DIs) can be defined as an extended continuum of the previous definitions of information technology (IT) applications and systems, as they refer to organizational or societal information infrastructures with functions that enable their existence, for example, as industries and enterprises (Tilson et al. 2010b, 748). This article draws from the most recent Information Systems Research (ISR) for defining DIs (e.g. Tilson et al. 2010a; 2010b; 2012). Infrastructures, in general, are thought as widespread structures, such as transportation or drainage networks that are used as taken-for-granted in everyday activities (Star and Ruhdeler 1996, 112-114).

DIs have been defined as ‘a shared, open (and unbounded), heterogeneous and evolving socio-technical system (which we call installed base) consisting of a set of IT capabilities and their user, operations and design communities’ (Hanseth and Lyttinen 2010, 4). DIs spread over specific systems’ boundaries and applications’ functionalities. Moreover, they are entities that are supported by global, national, regional, industrial, or corporate infrastructures, and their development is obvious at both global and local levels. In general, DIs are the basis for and the result of digital convergence i.e. the increasing use and availability of technologies (Tilson et al. 2010b, 748–750).

In addition to drawing on the recent conceptualization of DIs in IS research, this article is informed by the widely cited concept of information infrastructures deriving from STS research, particularly from the work of Star and Ruhdeler (1996, 112–114). In STS DIs are not perceived as ‘a thing’, rather they are understood as situated practises. DIs are described by a set of salient characteristics which are embeddedness, transparency, reach or scope, learned as part of membership, links with conventions of practice, embodiment of standards, built on an installed base and becomes visible upon breakdown. An infrastructure is configured of these relational properties when ‘-- the tension between local and global is resolved’ (Star and Ruhdeler 1996, 114), meaning that large-scale digital technologies can be used and modified for local practises by local practitioners. A meaningful aspect is also DIs' contextuality and time dependency. These characteristics relate to the relational nature of DIs, i.e. it is not only a concept of ‘what it is’, but also ‘when it is’ (Star and Ruhdeler 1996, 111–114).

3. Digital Infrastructures as Potential Research Possibilities for Archaeology – Past, Present and Future(s)

In the modern world the overall expansion of information technology has enabled the development of DIs, and simultaneously it has affected and changed social structures. The socio-technical ISR has studied the development of DIs. Technological devices, software, and applications became cheaper and capable of communicating, storing, and processing information. This process created device convergence in the same network, and also enabled gradually networks to convergence. As a result of this progress, also social infrastructures were reshaped and this caused, for example, industry and market convergence. This digitalized world created DIs and enhanced generativity, which means that individuals, groups, and organizations are able to co-create technology, services, and information, and further both transfer and use information and services in the networks. This has had an effect on, for example, service divergence such as wider participation and production in new services and emergence of
new market conditions. This has changed the previously more stable society and resulted in new forms of social and institutional order, and created questions, such as who controls the creation and offering of new information-based services. (Tilson et al. 2010b, 750). All these have an effect on existing DIs and their future developments. (Tilson et al. 2010b, 750). In that development, term ‘digitizing’ refers to the techniques of producing digitized documents, whereas ‘digitalization’ refers to a broader level socio-technical process where digitizing techniques are used (Tilson et al. 2010a, 3). This is the foundation for the evolving infrastructures (Tilson et al. 2010b, 749).

This progress goes on also in the cultural heritage management area, yet in a less business oriented mode. Increased use of technologies, the emergence of digital data, and the possibility of shared procedures in the future allow DIs to emerge and evolve. Technologies and networks may create possibilities for new research practises in archaeology, by data combination, and by making research collaboration and access to data possible today and in the future. Therefore, there is a need to reflect on and study these areas in also in archaeology as Kansa, Whitcher Kansa and Watrall (2011) have done. The major research challenge with DIs is that they are heterogeneous, complex, and relational in nature. To move beyond DIs and try to get a holistic understanding of their infrastructural nature and the factors that affect their successful existence, maintenance, or breakdown, is one of the major aims in research of DIs (Ribes and Lee 2010, 238–240).

For our knowledge, this kind of research does not exist in archaeology from the perspective of DIs, instead there have been attempts to collect recent results and experiences with technologies (such as Kansa et al. 2011). For this is a reason, infrastructures are still in the development stage in archaeology.

The research area of eScience and DIs is interdisciplinary and creates a wealth of possibilities for domain areas, such as archaeology. It offers viewpoints from the practice of general digitalization, to research in changing requirements and conventions in archaeology, especially in the use of technologies, and the lifecycles of developing DIs. Further, developing DIs could be analysed from a specific viewpoint as socially constructed, dynamic, and diverse structures using shared procedures and technologies, and in future, existing conceptual models (Tilson et al. 2010b, 755) for social dynamics could be tested for archaeological DIs. In archaeology, studies could be done on development and changing memory practices (Bowker 2005), for example standardisation, conceptualization and information models, or on associated reflections of research and interpretation processes in archaeology as case studies.

Studies in eScience have often focused on commonplace activities, developments, opening black boxes in DIs, and researching the lifecycles of DIs. Further evolution, sustainability, standardization, and integration of DIs are other general research topics in eScience. Ribes and Lee (2010, 239–242) proposed three key areas of study relating to DIs and their future design, which are also meaningful from an archaeological perspective: ‘(1) how to support domain specific practices, (2) how to facilitate data sharing, and (3) how to support virtual organizations that have a component of traditional organizations’. These are challenges to be faced in the evolution of DIs, but they are also relevant from an archaeological perspective – maybe not in the archaeology praxis just yet, but certainly in future (Ribes and Lee 2010, 233–242).

4. Socio-Technical Dynamics for Digital Infrastructures

Socio-Technical Dynamics are one important factor – partly emergent, partly explicit – affecting DIs. Tilson et al. (2010b, 755) have demonstrated the social dynamics of DIs from a very high-level perspective.
as a synthesized conceptual model based on literature analysis and previous studies. Recently Tilson et al. (2012) have also applied their conceptual model in a case study. In that, flexibility and stability are crucial to the process of evolving and developing DIs. DIs need to be flexible to be able to evolve, but they also need to have a stable basis enough, meaning that they need to be simultaneously controlled and autonomous in nature. That leads to a paradox of change and control in DIs, meaning, that technical stability is needed for their functionality, but a stable technical foundation allows for socio-technical connections and a social constancy that are needed for agreement between data definitions and user interfaces. The consequence of this process is standardization that allows DIs to grow. This can lead to changes in social structures and boundaries and to the widening of operational environments (Tilson et al. 2010b, 753–757).

Technology and standards can be analysed as control points, as mechanisms that control connections in the socio-technical system, such as domain specific practices, and use of products and applications. Standards are needed for stability, which means both loosing and increasing control. As a consequence, infrastructures may restrict the emergence of new infrastructures, and this leads to conflicts (or tussles) between new and existing control mechanisms. Moreover, they can be seen as actors or devices for actors, which can exert new control points or negate others, for example, they can exercise power. (Tilson et al. 2010b, 751–757). Recent articles have presented research on mechanisms for co-creating technology, services, and information, which could be used to produce guides for designers, managers, and policymakers (Tilson et al. 2010b, ibidem).

Archaeological data, information models and evolving standards (such as Binding, May, and Tudhope 2009) and technologies and their users with other stakeholders have a position in the social dynamics in DIs, and applying this kind of concepts to archaeology is one research perspective, which has not been done yet for evolving archaeological DIs. From an archaeological perspective, the existence of contesting control points may not be so obvious or extreme; it may be more moderate, such as differences in domain specific research traditions and practice and research methods and targets. However, the general tensions associated with the relational notion of infrastructures and their dimensions (Star and Ruhdeler 1996, 112–114), such as tensions between global and local development, apply in archaeology. On the other hand, generativity, control points, and strategic actions are also active operations undertaken by users (i.e. bottom-up approach) and authorities in institutions (i.e. top-down approach). Therefore, DIs are not just self-evolving but also consciously developed for archaeology, often with an agreed understanding that top-down created standards are not solution for archaeological standards (Dunn 2011, 97–117).

STS viewpoint offered by Edwards et al. (2007) presents an overall research agenda that is based on relating research on emerging cyberinfrastructures (Am.) to more historical studies of (information) infrastructures. The evolution of DIs is achieved by a technology transfer from one place or domain to another and it can be adapted to new areas. Infrastructures emerge when controlled local systems are linked to more heterogeneous networks, which need coordination and control. Entries, technologies, and standards are needed to link the heterogeneous networks together. The development of DIs is challenging and problems arise when it comes to stabilization, and later, when they need to be altered. Decisions made in the beginning are important in this process: the design of DIs should consider the future and allow for flexibility. Otherwise, this process can create tensions and ‘orphans’ of infrastructure that do not accommodate into the DIs. DIs can be described as having technical, cultural, legal, financial, and organizational variations and adjustments (Edwards et al. 2007, 6–8, 24–
Designing for DIs is compared to getting between social and technical infrastructures and is described as finding the tools for designers to match the true size of a solution and to trying to understand when something needs an organizational or technical fix (Edwards et al. 2007, 33). This viewpoint is based on Star and Ruhleder’s (1996) notion of infrastructure and uses their more elevated term, tension, to describe challenges in DIs.

These insights could help archaeologists to analyse the evolving archaeological DIs, increase awareness of their nature, including opening (i.e. visions of eScience) and restrictive effects (e.g. control points), as they relate to changing memory practices, such as the role of standardization and use of technologies in DIs. The research in other disciplines relating to DIs could be further utilized to anticipate DIs’ broad affects, possibilities, challenges, and controversies, which may lie ahead. The subject of this paper could be deepened to address specific questions relating to infrastructures, such as their design, control, power, and value creation in these wide-scope structures (see for example Hanseth and Lyytinen 2010b; Elaluf-Calderwood, Hertzoff, Sørensen and Eaton 2010; Law 1992). These questions could be applied and synthesized in archaeology. Control, conflict, and value creation aspects of DIs (Elaluf-Calderwood et al. 2010) could also be used to address global and local value creation in archaeology. For instance, the cultural heritage value as tangible and intangible values (United Nations Educational, Scientific and Cultural Commission 2012) is recognized but difficult to measure. Similarly, the economic value of the future uses and benefits of the data is difficult to specify. One example of this is the reuse of shared data (Lassila 2012). These kinds of uses and conceptualizations of DIs could make it easier for archaeologists to justify the need for archaeological DIs and get the support needed to create them.

To summarize, DI technologies, standardization, flexibility, control, and the visions of eScience are interrelated. Flexibility is needed to allow DIs adapt to future changes. The flexibility versus control aspect, i.e. the paradox of change and control, is tangled with questions about the practical design of DIs and theorization of their design (Hanseth and Lyytinen 2010). When the paradoxical and social nature of DIs is revealed, this can be analysed from technical and social viewpoints, which creates a wealth of research opportunities (Tilson et al. 2010b, 757). Another problem with the current eScience research is that methodological and theoretical work relating to DIs is lacking. Current approaches are only capable of making the complexities of DIs perceptible (Ribes and Lee 2010, 237).

5. Digital Infrastructures and Archaeology

In archaeology, the needs to increase global interoperability between diverse data sets and to guarantee future use of data have arisen. Relatedly, lack of shared standards and practises has been stated widely as a problem for achieving interoperability, although it is acknowledged simultaneously that nuances in local archaeological data should be preserved (for example Snow, Gahegan, Giles, Hirth, Milner, Mitra and Wang, 2006). Standards are used to achieve data interoperability in DIs (Ribes and Lee 2010, 235), however, standardization is not just the method used to attain the goal (i.e. interoperability), theoretical frameworks relating standardization processes also exist (Fomin, Keil, and Lyytinen 2003), which could facilitate the standardization process.

For the development of DIs the standardized information models, such as CIDOC object-oriented Conceptual Reference Model (CIDOC CRM), are essential as they form the ‘semantic glue’ component of DIs. CIDOC CRM offers an extensible semantic framework, a formal language that mediates information and data from disparate local data sets for global
use. It is widely used for digital cultural heritage data in order to increase interoperability between datasets in different disciplines. It has an emerging status as a shared ‘ontology of culture heritage information’, exemplified by the attained official standard certificate, ISO 21127:2006 (International Council of Museums 2012). While there exists no consensus about standards in archaeology, there is yet a serious effort to apply CIDOC CRM to archaeology (Binding et al. 2009).

Furthermore, digital curation, the combining of rich and diverse data sources produced by heterogeneous actors, and the management of data have proved challenging in archaeology (Kansa et al. 2011; Kintigh 2010; Snow et al. 2006; Richards 2002; Richards and Hardman 2007). Therefore, standards, including technical, metadata, and content standards, are needed (Richards 2009, 27–28). Standards can form a basis for the evolution of DIs. As control points, standards, and technologies can have implications for archaeological fieldwork practices, knowledge production, and scientific procedures, archaeological theoretical and methodological discussions reflect on these relationships as ‘going beyond the tool’ (Huggett 2012). However, this reflection does not have an awareness of the conceptualization of DIs or how tools can be seen as a part of evolving DIs. Using the expression of Tilson et al. (2010b), DIs are currently a ‘missing agenda’ in archaeology, but the need for them has been recognized (Huggett 2012).

It is possible to retrospectively trace how the need for digital access to archived archaeological data has developed along with the increasing numbers of technologies and possibilities for creating and using digital data. These developments have created a preliminary basis for archaeological DIs. Huggett and Ross (2004) wrote about the need for interoperability, transparency, and long-term access to archaeological information resources, as this would provide a more refined understanding of the implications of information uses and processes. This understanding should cover the whole process relating to creating, processing, and managing archaeological information, as continuum.

Further, according to Huggett and Ross (2004) archaeological analysis has firmly relied on hypotheses and the creation and testing of models. However, it has also been emphasized that interpretation has to go beyond these models and technologies in order to find motives, impacts, forces and logics behind large-scale systems i.e. why they were employed, what their future might be, and what the nature of change is. Today, these questions remain open, because change is continuous. The lack of a consistent interpretative approach has been acknowledged, as has the need to ‘open up black boxes’. ‘The black-boxing’ has been understood in social science oriented archaeology as technology use that does not consider humans in the process and uses as black-boxed interpretations in scientific research. (Latour 1981, cited in Jones 2002, 34–46). Similarly, Jones (2002, 34–46) refers to the invisible phases and decisions in work practices as ‘black-boxes’, for example, when processing and interpreting archaeological data. These discussions point out that interpretation in archaeology is about data and information (Huggett and Ross 2004). This has a crucial influence on the evolving archaeological DIs – there is a huge range of variables effecting on archaeological research practises and interpretation. That is a typical feature of DIs, as they are socio-technical in nature (Tilson et al. 2010b, 748–9).

There is also a demand for the archaeological DIs. In 2006, Kintigh reported the need for an archaeological information infrastructure ‘that will allow us to archive, access, integrate and mine disparate data sets’ (Kintigh 2006a, 567). He presented a vision for cyberinfrastructure (2006a, 568–578) and a schematic design for a knowledge-based archaeological data integration system (2006b,
Lately, he has written about sustainability of the archaeological digital record, its financial model, and problems faced in practical work (Kintigh and Altschul 2010). In addition, the ‘need for service-oriented’ archaeological cyberinfrastructures has been highlighted, in relation to electronic and user-friendly data archives in archaeology (Snow et al. 2006, 958–959). Most recently, the edited volume by Kansa et al. (2011) discussed archaeological cyberinfrastructures and collected the most recent studies, challenges, and possibilities of pioneering projects, such as Semantic Web Technologies and Linked data and other web-based technologies for data combination solutions.

The need for awareness of DI concepts has gradually been realized in archaeology, but the large amounts of different existing practices have been seen as a challenge for formulating the principles of DIs. The specification of DIs is challenging, as future purposes of research need to be integrated. Defining the proper requirements has been discussed recently on a general level, and qualitative data gathering has begun to collect these requirements from different domains in humanities, including archaeology (Benardou, Constantopoulos, Dallas and Gavrilis 2010).

Archaeological processes begin with field practices, like an excavation, and end with digital curation, and interpretation is an important aspect in every stage of archaeological work practices. To synthesize this process briefly, archaeological field practices and data processing are used to select and make sense of the past (Papaconstantinou 2009, 5) that is fragmented (Jones 2002, 40). Research, interpretation, and data related questions in these processes are often complex and technologically aided (see for example Huggett and Ross 2004; Lock 2003, 7; Lock and Molyneaux 2006; Richards 2002, 2008, 2009). According to Witmore (2004, 156–159; 2009, 515–525, 532) fieldwork practice and data processing can be described as a memory practice, which consists of a data translation that includes several steps to reiterate and transform data and tries to anticipate future demands relating to data. This practice has been defined as ‘the design of the information worlds for future’ (Witmore 2009, 532). Archaeological processes also include ‘black boxing’ of data but also opening of those ‘black boxes’ (Jones 2002, 34–46) when reporting data and delivering it as a report. In that process, collected archaeological data is transformed into cultural heritage data (Papaconstantinou 2009, 1). In the process, the data is finalized in a digitally curated form (Higgins 2008) but also continues to have a future use.

With a specific angle on the lifecycle of archaeological data, it can be observed that data collecting and processing form the basis for and intertwine with all interpretation and translation processes of subsequent research. Archaeological field practices and data processing have been analysed from numerous points of interest, such as general work practices, research or documenting processes, interpretation processes, preservation processes, and cultural heritage data creation processes. The formulation of requirements for multi-discipline DIs (such as humanities including archaeology) is challenging, because they should cover all the important aspects derived from each discipline (Benardou et al. 2010).

The challenge for the development of archaeological DIs is to understand these practices and integrate eScience visions with the archaeological visions. Another challenge is dealing with the possible tensions between the needs of local and global actors in the network of social dynamics as they evolve in archaeology or cultural heritage management. In addition, there are actors in various stages of digitalization development in the world, some of whom are incapable of taking part in this process. Awareness of the importance of standardization and future requirements relating to data gathering and processing
should be increased, but the local peculiarities should also be preserved. Searching for balance between these different goals is a long-term process.

Finally, Huggett (2012, 212–213) concluded his recent paper by stating that archaeologists are able to think about their tools, but going beyond them and to research implications of technologies, is more difficult. He said in trying to reach ‘the beyond’ researchers may lose sight of the archaeological research target itself. Thus, it is important to be aware of these questions but not to be unreflectively drawn into them (Huggett 2012, 212–213). DIs and related theoretical constructs that exist in socio-technical ISR and STS research offer approaches for ‘going beyond the tool’, and help addressing ‘the beyond’ as complex socio-technical phenomena in contemporary and future heterogeneous digitalized networks. These definitions are applicable in developing DIs in the archaeology or cultural heritage management field. The evolution of DIs connects archaeology to similar global processes, which are on going in other digital data driven or technology-oriented domains.

6. Conclusions

The aim of this article was to introduce the concept of DIs and the associated model of socio-technical dynamics and relate them to archaeology. The article proposes the following points, which the conceptualization of DIs together with the notion of Socio-Technical Dynamics offers archaeology in the eScience era:

- A new interdisciplinary position and context to analyse and theorize the methods and practices used in archaeology.
- Further awareness of evolving DIs and an increased understanding of the involved socio-technical dynamics (e.g. the meaning and effects of control points) could deepen these interdisciplinary reflections for archaeology.
- Archaeology can also learn from the experiences of other disciplines and anticipate DIs’ broad effects, such as possibilities, challenges, and controversies in order to prepare for the future.

With the use of theoretical frameworks and models derived from other disciplines, archaeology can be analysed with a fresh viewpoint that may allow for new insights. We hope to have illustrated that the concept of DI offers a valuable theoretical model for both addressing practical problems and considering research questions in eScience era archaeology. A sound theoretical framework may contribute to not only archaeological research but may also create a more generic model that will fulfil the needs of designers, managers, and policymakers (Tilson et al. 2010b, 751, 753).

References


Geospatial Technologies and Analysis
Intrasite Spatial Analysis of the Cemeteries with Dispersed Cremation Burials

Marge Konsa
University of Tartu, Estonia

Abstract:
The aim of this study is to find appropriate methods for the analysis of dispersed cremation burials in order to understand the funerary rituals and the formation of cemeteries. The point pattern data from the 10th–13th century AD cemetery at Madi in Estonia were analysed using several intrasite spatial and geostatistical methods such as autocorrelation, nearest neighbour analysis, point density analysis, and minimum distance analysis. As a result, a new understanding of the funerary rituals performed on the burial place and formation process of the dispersed cremation cemetery is presented.

Keywords:
GIS, Intrasite Spatial Analysis, Point Pattern Analysis, Cremation Burials

1. Introduction

The essential characteristic of dispersed cremation burials is their collective nature and intentional indistinguishability of the individual burials. Burned bones and artefacts are usually scattered over the whole cemetery area in an irregular manner, although some cremation burials may have been put into pit as well. Most of the cremations are not placed in any kind of a container. The intrasite spatial and quantitative methods which are commonly applied for cemeteries with separate individual burials are not suitable for the analysis of dispersed cremations because of the specific nature of the latter.

Cemeteries with dispersed cremation burials were one of the main burial type in Estonia during the Iron Age (500 BC – AD 1225). The structure of the cemeteries varies a lot. Different types of above-ground stone constructions are represented, as well as burials under level ground. Previous research has been mainly concentrated on dating of grave goods and regional overviews of morphology or distribution of the cemeteries (e.g., Laul 2001; Mandel 2003). Intrasite analyses of burial places in Estonia have been made manually relying on simple visual inspection and restricted range of data. Attempts (e.g. Mägi 2002) to distinguish individual burials according to “gendered” artefacts have been unsuccessful because of unreliable methodology.

The aim of this study is to find appropriate methods for the analysis of dispersed cremation burials in order to understand the funerary rituals and the formation of cemeteries.

2. Madi Cemetery: a Case Study

This research is based on the case study of the 10th–13th century AD cremation cemetery at Madi in Estonia. The cemetery is characterized by the extensive area covered by sparsely located stones which were only weakly visible above the ground. The total area of the cemetery is 1600 square meters of which 65% is excavated. Excavations in Madi were carried out in the 1920s and 1960s, and results remained unpublished. However, the excavations in the 1960s (977m²) were well recorded and can be used for spatial analysis (Figs 1-2). On the excavation plans (scale 1:25) the distribution of artefacts, bones and charcoal was marked with dots that allowed to digitize them as a point data. There was altogether 2477 findspots of artefacts (including 1225 findspots of potsherds), 1603 findspots of bones and 1242 findspots of charcoal (Fig. 3). Amongst the grave goods were iron weapons, tools and
Intrasite Spatial Analysis of the Cemeteries with Dispersed Cremation Burials
Marge Konsa

utensils, as well as jewellery of bronze, metal parts of clothing, horse harnesses and riding gear. Most of the artefacts have been fired, and some of them were deliberately damaged (ritually killed). Bones are highly fragmented and calcined. In addition to humans, animal bones are also represented (Engbring 2011). Even though the distribution of bones and charcoal was recorded as point data during excavation, they were collected by 1-m squares.

3. Research Methods

Spatial data of artefacts, bones and charcoal were studied on two levels. First, the material from the cemetery was analysed as point pattern data with the aim to find regularities in spatial distribution. After that, smaller clusters and concentration areas of finds and bones were taken into closer consideration. Different analysis methods were used: autocorrelation, nearest neighbour analysis, point density analysis, and minimum distance analysis. Most of them are historically important methods in spatial archaeology, and they have not lost their value even today (Hodder and Orton 1976; Conolly and Lake 2006).

4. Results of the Intrasite Spatial Analysis

One of the primary methods of point pattern analysis is tests for randomness. The most proven statistic for this is the nearest neighbour index. It compares the distances between the nearest points with the distances that would be expected on the basis of chance (Ripley 1981). According to the nearest neighbour statistic the distribution of artefacts, bones, and charcoal in Madi cemetery was clustered (Table 1).

Density analysis allows one to describe and visualise the changing frequency of observations that occur within a given area,
often to compare different phenomena within the same area or against the same phenomenon in different areas (Conolly and Lake 2006, 173). As a visualisation method the point density with kernel density estimation gives a smooth and easily interpreted continuous surface for cluster identification (op cit, 177). The kernel density maps (Fig. 4) enable comparisons between the locations of clusters of artefacts, bones and charcoal in Madi cemetery. The general distribution area of bones was considerably smaller then those of artefacts and charcoal. Bone finds were situated closer to each other compared with artefacts or charcoal. The mean nearest neighbour distance between bone findspots was 8 cm, between charcoal and artefact findspots respectively 11 cm and 23 cm (table 1). The correlation between distribution of artefacts and charcoal (r = 0.5) was moderate. A few high density clusters of artefacts and charcoal were overlapping with each other. There was a partial overlap also in the location of high density clusters of bones and charcoal.

The most striking discovery, however, is the difference in distribution pattern of bones and artefacts. The correlation between general distribution of bones and artefacts was very weak (r = 0.1). Their distribution areas of high intensity clusters don’t overlap with each other. Described results reveal that the cremated bones have been distinguished from the rest of the pyre remains and treated separately from grave goods. This inference does not support the interpretation of funerary rituals expressed in previous studies according to which individual burials could be identified based on distribution of the artefacts (Mägi 2002). The northwest-southeast orientation of the high intensity clusters area of bones in Madi coincides with the general orientation of the cemetery (Fig. 1). Thus, the bones were placed mainly in the central part of the cemetery while the highest concentration of grave goods was along the edge areas.

Autocorrelation Moran’s I statistic was used to examine the spatial formation of the cemetery at Madi. Spatial autocorrelation indices identify whether point locations are spatially related. If neighbouring locations tend to have similar values to each other, then the spatial variable is said to exhibit positive spatial autocorrelation. On the other hand, if they tend to be different, then the variable is said to exhibit negative spatial autocorrelation (Wheatley and Gillings 2002, 118). At Madi cemetery there was a negative autocorrelation in the spatial distribution of well-dated artefacts. It could mean that the majority of the cemetery area was in constant use and grave goods from different time periods were lying next to each other.

Spatial co-occurrence of artefact types was studied with minimum distance analysis. With the help of the distance calculator tool the pairs of objects with closest distance were identified. Within distance² up to 20 cm the most frequently occurring combination was pairs of

<table>
<thead>
<tr>
<th>Data set</th>
<th>Sample size</th>
<th>Mean nearest neighbour distance (m)</th>
<th>Mean random distance (m)</th>
<th>Nearest neighbour index (R)</th>
<th>Z – test statistic</th>
<th>Distribution pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artefacts</td>
<td>2477</td>
<td>0.23</td>
<td>0.34</td>
<td>0.6766</td>
<td>-30.7958</td>
<td>clustered</td>
</tr>
<tr>
<td>Bones</td>
<td>1603</td>
<td>0.08</td>
<td>0.37</td>
<td>0.2151</td>
<td>-60.1227</td>
<td>clustered</td>
</tr>
<tr>
<td>Charcoal</td>
<td>1242</td>
<td>0.11</td>
<td>0.48</td>
<td>0.2214</td>
<td>-52.4930</td>
<td>clustered</td>
</tr>
</tbody>
</table>

Table 1. The result of the nearest neighbour analysis.

Figure 4. Kernel density estimates of the artefact (1), bone (2) and charcoal (3) distribution in Madi cemetery. Diameter of 2-m radius.
Intrasite Spatial Analysis of the Cemeteries with Dispersed Cremation Burials

Marge Konsa

different types of jewellery (fragments of breast chains, bracelets, neck rings, decorative pins, and brooches), as well as a combination of a knife and jewellery. These items were probably together on the same pyre and were placed afterwards into the burial ground as one unit. Similar combinations of artefacts are common grave goods in inhumation burials from the same time period.

The distribution pattern of different artefact types was studied with nearest neighbour statistic. Their distribution was mainly clustered or random. Within this general pattern there were only two exceptions – the distribution of knives and strike-a-lights was regular. Altogether 77 knives and 22 strike-a-lights have been found in Madi. The high number of knives is not surprising as they are the most common grave-goods found in 10th–13th century AD burials in Estonia, while a large quantity of strike-a-lights in Madi is exceptional compared to the other cemeteries. The phenomenon of regularity in distribution of grave goods is difficult to explain.

With smaller data sets, it was possible to draw conclusions based on simple visual inspection. For example the location of the dog bones and bear claws is noteworthy (Fig. 5). They were distributed only in a certain part of the cemetery. Seven bear claws out of ten were found in the same cluster with dog bones. It is likely that the bear claws could originate from bear-skin shrouds as there are no other bear skeletal elements present (Engbring 2011; Schönfelder 1994). Dog bones are found in many Estonian cremation cemeteries, while bear claws are rare (Maldre 2003).

The total number of burials must unfortunately remain open. There were a minimum of seven individuals buried, based on the osteological analysis (Engbring 2011). It is a very small number compared to the amount of grave goods found in the cemetery. It is possible that the missing bones have been deposited elsewhere, or some artefacts could be additional offerings besides grave goods, although the amount of artefacts which have not been in the fire is modest.

5. Conclusions

There is a long tradition of successful applications of point pattern analysis in settlement site studies. This paper was the first attempt to use similar methods in the research of the cemeteries with dispersed cremation burials.

The results of the case study show that the distribution of artefacts, bones, and charcoal was clustered in the cemetery at Madi. The correlation between general distribution of bones and artefacts was very weak. Apparently the cremated bones have been distinguished from the rest of the pyre remains and treated separately from grave goods. The bones were placed mainly in the central part of the cemetery while the highest concentration of grave goods was along the edge areas. There was a negative autocorrelation in the spatial distribution of well-dated artefacts, which could mean that the majority of the cemetery area was in constant use.
use over two centuries. Artefacts with closest distance were different types of jewellery, and combination of a knife and jewellery. These items were probably together on the same pyre and were placed afterwards into the burial ground as one unit.

Thus, with the help of intrasite analysis it was possible to identify general patterns in the treatment of cremains (i.e. pyre remains), as well as to find out the variability in burial practices. As a result, we now have a better understanding of the formation of this type of cemetery and more information of the funerary rituals performed at this burial place. The methods used for the analysis of the cemetery at Madi can be applied to other similar sites as well.

Acknowledgements

This research was supported by the European Union through the European Regional Development Fund (Centre of Excellence of Cultural Theory) and by European Social Fund’s Doctoral Studies and Internationalisation Programme DoRa. I would like to thank from Southampton University Dr. David Wheatley for his advice and Laurel Engbring for carrying out the osteological analysis.

References


A Specific Approach for a Peculiar Site: New Spatial Technologies for Recording and Analysing a Palaeolithic Site (the Cave of La Garma, Northern Spain)

Alfredo Maximiano, Pablo Arias and Roberto Ontañón
Universidad de Cantabria, Spain

Abstract:
The archaeological site of La Garma (Cantabrian region, Spain) is one of the most impressive Upper Palaeolithic sites in Europe. Part of a complex karst system, the cave contains around 800 square metres of exceptionally preserved Magdalenian floors, including ca. 76,000 archaeological items (bones, flint and antler implements, marine mollusc shells, portable art objects...). At least six structures can be directly observed, without any excavation. Yet the site is also a very challenging archaeological document. In these circumstances, we are close to mapping and quantifying peculiar spatial distributions (material consequences of a set of social actions) and not only a fractional spatial distribution subjected to various alteration (natural and atrophic) processes. We might then elaborate consistent solutions where dispersions of material effects and their causal actions configure a relational and interpretative discourse about spatial variance. This situation is an ideal laboratory to study not only multiple, different and peculiar spatial distributions of diverse collections remains. It is also an ideal environment for i) applying new spatial technology (capturing process and management data); ii) working with new approaches in spatial analysis (geostatistic and analytical visualization); and iii) re-thinking classical issues of spatial analysis in archaeological contexts.

Keywords:
Magdalenian Floors, Data Capture Process, Spatial Analysis

1. Introduction

The archaeological site of La Garma is situated in the north of Spain, near the city of Santander (ED50 geographic coordinate: 43º 26´ 1.79 N; 3º 39´ 52.72 W) (Fig. 1). This cave is integrated in a complex karstic system where the original access collapsed during the late Glacial period. The entrance to the gallery was blocked and this interrupted geological and anthropogenic processes. This particular circumstance has allowed high levels of preservation on the Magdalenian floors (Fig. 2). Today, we can detect an extremely high density of items with a certain consistency in the occupation floors inside the site. Around 76,000 archaeological objects (bones, flint and antler implements, marine mollusc shells, portable art objects...) are spread on the floor of the cave.

In La Garma, the main goal of archaeological spatial analysis –reconstructing the activity performed in the past in specific places from the partial distribution of remains– acquires a new dimension, possibility and opportunities. Because of the particular conditions of the site, the archaeological record is close to the last set of social actions performed in every “cluster” of material localizations. It is a possibility to map a distributional plot of actions, beyond the classical spatial study where the main object is mapping spatial distributions of remains in term of their nominal categories, and then try to identify the social use of space.

For an adequate interpretation of spatial variance, it is necessary to establish a balanced discourse between materiality record (partial consequences of actions), possible causal action (the origin of material effects and their
distribution) and quantify spatial structure of remains (spatial analysis). In this complexity, we must think of an appropriate environment to integrate data already collected in the last fifteen years with determinate technology (general topography of the cave, archaeological work in Zone IV...) and establish the integration of these data with new spatial technologies which are implementable to capture and manage data in other areas of the cave.

On the other hand, we must note that La Garma is a very complex site in terms of documentation of its material evidence. This is because there is duality in the archaeological level activity and the conservation strategies. This site was included in the UNESCO World Heritage list in 2008 and under this circumstance it is necessary to conciliate the measurement of scientific record with a coherent strategy in preservation and integrity of its evidences.

The evidence inside the cave displays a large variability of material effects (spatial structuring) related to certain social actions. An interesting example of this relation is found in the link between parts of the occupation floors (processing of pigments) and certain examples of cave art (Arias et al. 2011). As there is a significant repetition of remains in/with a series of locations, the presence of possible structures, parietal art..., we consider that La Garma complies with a series of requirements and therefore offers the optimum conditions to be able to approach the distribution of the last actions carried out in the cave, before the entrance collapsed.

Under these circumstances, this paper presents a proposal about the coherent use of spatial technology in data collection and management, for the development of an appropriate research environment. With this perspective we aim to test a series of hypotheses about segregation and possible use of space, divided into two main units: areas used for activities connected with production and consumption, and others connected with the symbolic or ritual realms.

2. How can we Capture and Manage (Efficiently) Archaeological Events at La Garma?

Great advances in computing and data management in archaeology have been made since the 1990s. Today, the protocols for the capture of archaeological spatial information still follow, more or less, standards of digital format (bits) with different means to capture
them (total station, laser scanner, differential GPS and photogrammetry) where the most coherent way to manage these data is a GIS platform in which a spatial data infrastructure can be developed, adapted to the needs of each project. The case of La Garma presents ideal conditions for data capture and management due to the large population of remains (approximately 76,000 items) and the conditions of the cave itself.

Working with a GIS platform will enable interactivity with the data as it will be possible to obtain, among others: 1) connectivity with all types of geodata capture systems (qualitative data, total station, photogrammetry, laser scanner, etc.). 2) conversions of spatial entities depending on the needs. For example, to generate products aimed at heritage diffusion, generate specific utilities for analytical series (e.g. algorithms for texturization of polygonal entities; generate points clouds for certain analytical series; recognize shapes, edges and boundaries of certain items for morphometric analysis... 3) Manipulation of components for restitution and spatial modelling... 4) hybrid environment where real and virtual entities can be interaction.

In this sense, the advantages of a GIS platform for data management at this archaeological site can be summarized in the following points:

- Easy method to design spatial relations (topology) (Despite low capacity at Computer Assisted Design).
- User-friendly spatial analysis outputs (point patter, neighboring, clusters, semivariance...) and High and Complex computing power.
- It is a multidimensional model of archaeological events. It is more than a drawing, it is a analytical and visualization tool.
- Spatial and non spatial data can be analyzed simultaneously in a rational and relational way. (e.g. spatial distribution remains and rock art).
- Digital environment for virtual applications and hybrid environments: simulation process, spatial prediction; re-excavation in digital laboratory; analytical visualization.

However, we have chosen the type of platform to be used with our geodata, but the capture system is not so clear as regards the conditions/particularities of each of the possible systems in connection with the peculiarities of the site.

At La Garma, total stations have been used intensively, and thanks to this system topographic survey data have been compiled for the whole cave and groups of entities have been georeferenced (Rolleimetric system) (in Zone IV). From this experience, it is considered necessary to adopt a new system to record the totality of the site, as the investment in labour and the results obtained would be greater with the incorporation of other techniques in the data capture process. The alternatives being considered are the use of a laser scanner and/or photogrammetry. Both options are interesting, as the data compilation by either of these systems is enormous.

Applying these technologies the resolution levels and the data volume will be very high, and in this sense, we should not forget that these capture systems need intense post-process work as the information gathered is only a georeferenced casing. With this first product, calculation algorithms must be applied to be able to define the boundaries and discontinuities of each of the items (edge detection), contained within the casing. At La Garma, due to its peculiarities, this may need the use of a super-computing environment.

Once all the elements have been differentiated, and if we aim to achieve a valid database for the spatial analysis of distributions,
we need to synthesize the information further, convert the polygons into points (centroids) and accept a determinate degree of distortion. With current technology and workstations, it is likely that the investment of time and effort would be focused on post-processing of data rather than in the capture of the surfaces. Under these parameters, we believe that the technology used at the site to capture reality should be more in consonance with the accessibility in obtaining new information (i.e. centroids) from the original data than with the level of detail. For a balanced choice, it will be necessary to weigh up numerous aspects, like the level of precision of the data, the resolution needed (going beyond the simple idea of maximum resolution) for the objects represented, which takes us to a definition of the data quality. To this we must add the compatibility and interchange between different data formats depending on the technique used in their detection and measurement. In addition, we must be aware of the hardware resources required for the post-processing, as we will work with eminently visual scenarios (graphic output) as well as clearly quantitative environments. If we are able to resolve these questions, we will obtain efficient and highly effective information capture with which to achieve the objectives of the study.

3. Middle Magdalenian Floors at La Garma: Challenges for Efficient Spatial Analysis

The casuistic of spatially distributed events usually correspond to a change in the conditions of the system, to change in neighbouring distributions that exercise influence over each other (relational model) or instead, a combination of both effects. According to this, spatial analysis in any discipline aims, first of all, to characterise the observed distribution. That is, to know whether it is random or not, what its distribution model is and what the module of spatial correlation/dependence is. All this is necessary for the main objective in this kind of analysis: determine the reasons why the distribution is as it is, and predict the changes that might occur under certain conditions.

If we transfer this idea to archaeological cases, it will be through an exhaustive analysis of spatial variability that it will be possible to detect relationships between different spatially distributed events, such as certain material remains (knapping waste, fauna and so on), structural elements (like grouping of some kind of large boulders and the presence of hearths), or the distribution of chemical residues associated with a particular activity (like tanning hides and processing pigments). If we are capable of quantifying the relationship between events (i.e. the increase in the number of remains of type \( a \) in zone \( x \), \( y \) coincides with a decrease in types \( c \), \( d \) and \( f \) in the same area), we shall be able to make interpretations about some of the activities performed in certain places, and in this way go beyond the mere description of spatially distributed remains in a certain zone.
For this approach to be effective, we should be capable of going beyond the nominal categories (i.e. burnt bone, rim of medium-sized container, lithic knapping waste, and so on) for which archaeological spatial analysis usually holds little meaning. What meaning can it have if the distribution of wild fauna bone > 0.02cm³ tends slightly to spatial uniformity and is correlated at about 1.2m, whereas the distribution of wild fauna bone < 0.02cm³ is spatially concentrated intensely with a spatial correlation of 0.5m? This type of detailed analysis (spatial structure and correlation) would not provide much more information if we made a mere visual inspection of the distributions. On the contrary, if we reformulate the question into that of the nominal category being centred in the type of relationship that might exist between size and location, using the same values as before, we can propose the hypothesis that the size conditions the mobility of the remains. Therefore, this is a type of action (anthropic or natural) that discriminates according to certain parameters. Thus, it would make sense to parameterise the structure and the spatial correlation of each of the subsets.

This theoretical example is in many ways linear and simplistic, as it does not consider variables of the taphonomic kind, the archaeological activity itself, or whether it is an equipotential surface, but it is an example that may be quite comparable with the Magdalenian occupation floors at La Garma, due to the peculiarities of the site. That is, it is possible to detect and control the elements that distort/condition the distribution, because the distortion of archaeological activity itself does not exist (since it is not necessary to excavate to reach the archaeological record). In other words, the occupation floors are perceived in their original position and the taphonomy is slight and easily manageable due the conditions existing at the site in the last 15,000 years. However, we should be cautious, as in numerous archaeological spatial analyses a fictitious correlation is often established between location of remains and working areas. The problem with this argument is that the location of these remains is not necessarily the emplacement of the type of action usually associated with them. Examples of this kind of shift are found in numerous cases of ethnarchaeological literature documenting clear differences in correlating actions and the location of the material distribution of its effects (see examples in Politis and Saunders 2002, among others).

The current state of the work at the site allows a prototype of spatial analysis to be undertaken from geostatistic approach where spatial variance, polynomial surface and spatial gradient will be the principal analytical components.

The area proposed for this study is Zone IV, where fieldwork has provided a large dataset and information generated through different analytical series (topography, phytoliths, fauna, lithic remains, presence and distribution of chemical residues). At the present time, we are generating a new geo-database where from geo-referenced maps (CAD format), providing spatial information about individual positioned objects.

In this geo-database, each of the items possesses x, y and z coordinates and is described according to a wide range of variables (form, texture, composition and function) which enables a complete spatial analysis at both univariate and multivariate levels. With this data, certain statistical operators for spatial analysis, and the GIS platform as a multilayer analytical visualization tool (i.e. Craing et al. 2006), we can establish a spatial analytical series combining: i) visualization elements (i.e. kernel density, 3D histogram, polynomial surface), ii) structural analysis (e.g. NNA, Ripley’s K, Moran’s I, Mardian Correlogram), iii) quantification of the spatial correlation of each of the distributions (semivariogram) and iv) interpreting the space in key of attraction-repulsion action in determinate locations.
4. Final Considerations: New Possibilities for Understanding Archaeological Spatial Complexity

The site of La Garma offers numerous opportunities, but at a spatial analysis level the most important issue is the possibility of carrying out a full complex analytical series for determining how the use of space inside the cave was structured and managed. If we commonly define both the use of space and the distributions of archaeological remains in terms of the material and formal qualities characterising them, in the case of La Garma we can go further and establish a spatial analysis aiming to map actions in space, rather than distributions of effects. This is able to know structural limits (in/ out of housing units), intersections (some actions distort the effects of others) and lastly, operational discontinuities can define spatial dynamics through the localized portion of recoverable material evidences.

At this site, the “proximity” existing between the ensemble of material remains and the social actions that produced them creates an opportunity not only to know (almost as if it were an x-ray) certain practices by the social groups who occupied the site, but also enables a profound reflection on the methods of collection and processing data in spatial analysis, as well as reformulating the interpretation we might make about the spatial variability. With all this, archaeological spatial analysis will acquire a new dimension, which will probably have an impact on trends in archaeological spatial analysis at an intra-site scale.

The application of these analytical components is important in the case of La Garma, but even more significantly it offers an opportunity to assess the interpretative capacity of these tools, which are not designed for archaeology, for certain archaeological problems.

Kanevski (2008, 5-8) describes four phases in the analysis of spatial data: exploration of spatial data (ESDA); monitoring and description of the populations; structural analysis; and spatial-temporal prediction/simulation. These series are usually applied in archaeological cases at the exploration and description level, but proposals are rarely made about structure and prediction, for two main reasons: because of the particularity of some contexts of spatial data in archaeology and, above all, because the archaeologists carrying out the analysis have not been trained appropriately.

When the exploration, description and structure of each of the distributions has been made, all this information will enable the development of analytical visualization in a multilayer environment as we are interested in detecting and quantifying the possible relationships (whether in a certain set of locations, the increase in some variables produces an increase in others, a decrease or a lack of interaction) that might exist between different distributions. In this way, an approach is made to possible spatial management (for example, near a hearth, knapping waste is found in larger quantities and consumption remains are scarcer). This analysis may be numerical (multivariate), but if we select a multilayer graphic output where each of the layers is associated with the distribution function of each population, we shall obtain an analytical visualization (probability of distribution through a distribution function) where we can appreciate positions, forms and intensity of the population and its interaction with other populations. An interesting solution is multidimensional visualization in terms of colour channels (Craig 2006), where the intensity of each colour and the resulting combination of colours allows the visualization of spatial differentiation.

If the distributions are well-characterised, we shall be able to make a simulation of the causal process that originated them. This is possible at La Garma as the relationship between the distribution of remains and
the action(s) that might have caused this distribution is quite intense. If we determine the spatial trend of the causal action (for example, knapping action) we shall be able to estimate that a series of populations characterised by certain materials and function (knapping waste), which are spatially distributed in a certain way (e.g. truncated bivariate normal distribution) correspond to a certain type of action (production of lithic tools) in that place (Maximiano 2012, in press). In this way, we have taken a qualitative step forward in our study, as we are not mapping simple distributions characterised in terms of form, function and material, but are able to map the probable position of actions.

The complexity of this kind of simulations does not lie in the repetition of possible solutions (all those that a Montecarlo process can award statistical significance), but in the absence of a “catalogue” where different types of spatial distribution are related, with material, formal or functional conditions, and with possible causal actions. It is also true that such a catalogue should be understood as an eminently heuristic solution, completely remote from reductionism and linearity in the synthesis of the complexity in the cause-effect sequence that includes the fraction of the spatially distributed archaeological remains.

However, we can (and should) widen our perspectives in understanding what we aim to achieve with our spatial analysis. We should begin to overcome the tendency to use simplistic reasoning to interpret distributions. Not all the random distributions are distortion nor all the concentrations are the places where the action is articulated and thereby organises space. Therefore, the performance of simulations in our spatial analysis is recommendable, as they will provide a fuller vision of the distributions and their possible causes.

If we know the distribution structure, the distribution function characterising it and can associate the type of causal action, we shall be able to estimate the spatial rate of change of this distribution, and therefore define and calculate the use of space in terms of the locations that attract/repel the performance of actions. For example, if what we understand as a domestic unit (“hut”) is usually related with alignments of stones, post-holes, hearths and certain material remains, we shall be able to reformulate this type of emplacement, or rather, use of space, according to a series of convergence/dissuasion of actions on a series of locations. In this way, the concept of the use of space would acquire a new dimension and would be measurable according to a spatial function called spatial gradient (Maximiano 2008).

All this reasoning is applicable to the case of La Garma, which is a magnificent testbed. For example, the structures that have been documented (stone enclosures) contain within them a distribution of remains (special attention to the presence/absence of chemical residues) with certain characteristics that are apparently different from both the remains documented outside these structures and their distribution type. This enables the definition of domestic unit (use of space) based on fuzzy boundaries where the attraction/repulsion of actions manages the space. By calculating the gradient (derived from the distribution function) we can quantify these uses heuristically. We can thus overcome the classical concept in spatial archaeology in which the emplacement (concentration or pseudo-concentration) of remains is usually understood to be an indicator of the use of space.

5. Some Reflections

Coherence in spatial analysis using current technological tools may carry archaeological spatial studies to a new dimension, but it should be accompanied by an intense theoretical reflection (Barceló 2005). We cannot apply computational methods alone, as the answers to our problems do not lie in the use of technology; this is only mediation.
We understand that reporting the methodological implementation allowed by the archaeological record at La Garma may be a significant contribution to archaeological spatial analysis, both in its methodology and in its ontology.

There is no doubt that the solution does not lie in technological prowess or supposed analytical complexity, but in knowing how to use these tools and new ways of tackling spatial variability problems with coherence. We are aware that the possibilities of La Garma go beyond the comments made here, as we have not examined potential opportunities for this problematical like fuzzy logic, fractal dimension or, for example, formulated our spatial problem as a statement within the Theory of Evidence (Dempster 1967; Shafer 1976), to which archaeological spatial uncertainty is very well adapted.

However, it is equally true that before new aspects are tackled, we should explore and validate questions that one way or another archaeological spatial analysis has not been able to or not wanted to deal with. La Garma will allow us to approach certain aspects of the ways of life in the past, but will also generate a new debate about the possibilities of archaeological spatial analysis and new approaches to be tested.

References


582
Donald Institute. Cambridge University.

Use of Quantitative Methods to Study an Alpine Rock Art Site: the Mont Bego Region

Thomas Huet
Université Nice Sophia-Antipolis, France

Abstract: The Mont Bego site (Alpes-Maritimes, France), located 2,000 m above sea level, has more than 20,000 figurative pecked petroglyphs engraved across an area of 900 hectares. The site has been the subject of study over the last 45 years by the H. de Lumley research team. They collected the positions of 4,000 carved rocks and made drawings of thousands of carvings. After a 3-year collaboration with Lumley’s team, we developed a GIS, a database and protocols for registration and analysis of the carvings. The statistical analyses, for both the carved rocks and the carvings, included multivariate analyses, regressions and non-parametric tests. The crosscheck of spatial, parametric (such carvings coding) and iconographic data has permitted us to find regularities. All theses regularities (cost-distance from lakes, slope of carving surfaces, correlation of different carved themes, etc.) are measured quantitatively, forming a non-interpretative base of knowledge. These “archaeological facts” frame interpretations about the meaning of carvings. Concurrently with geostatistical analysis, an inventory and a revision of surimpositions of carvings has been realized.

Keywords: Rock Art, Prehistory, GIS, Statistics

1. Introduction

As noted by C. Alexander (2011), the study of the two major concentrations of rock-art carvings in Western Europe, Valcamonica-Valtellina (Lombardy, Italy) and Mont Bego (Maritime Alps, France), is still culture-historical and widely image-based, while databases, statistical procedures, and GIS analysis are little-used. For the first time in the history of Mont Bego research, systematic use of quantitative methods allows us to relate and analyze concurrently iconographic, parametric and spatial registers: a drawing of a carving is a piece of iconographic data, coding of this carving in a database creates parametric data which, when it is projected with a GIS, becomes spatial data. We will present here the majority of the different quantitative methods we used, along with some results.

2. Mont Bego

The Mont Bego region is in the French southern Alps, in the Mercantour National Park, 40 km from the Mediterranean sea. The site can be divided into two main sectors: “Les Merveilles”, with more schematic figures, and “Fontanalba”, with more naturalistic ones (Fig. 1).

Thanks to H. de Lumley’s team’s work - an inventory of almost all the carvings made over the past 45 years - it is possible to have precise knowledge about the importance of the site. Unusually in archaeology, the corpus of data is not a sample but rather the universe of remaining figures – although some may have been lost to erosion over time and new ones are regularly discovered during annual field work.

Mont Bego contains more than 4,000 carved rocks in a 900 ha area. It has some 20,000 figurative pecked carvings which can be divided into about fifteen themes. Based on a comparison between carvings of weapons and archaeologically-recovered objects, the carvings were first attributed to the Copper Age and Early Bronze Age (ca. 3300-1800 cal BC). Today, considering superimpositions, seriation, the spatial distribution of the different themes,
and some new archaeological finds, this chronology has been reappraised (Huet 2012). One must not forget that Mont Bego’s region is one of the first Alpine areas to be occupied during the Early Neolithic, as evidenced by the presence of Cardial ceramics (ca. 5500 cal BC), and was also occupied during Middle Neolithic (ca. 4250-3300 cal BC).

While the site has been known to archaeologists since the late 19th century, and continuously studied since the early 20th century, the meaning of these carvings is still discussed and different hypotheses have been advanced to explain their presence.

3. Hypotheses and Interpretations

Interpretations can be divided into two classes: “relational theories” and “non-relational theories”. Quantitative measurement and the use of hypothesis tests allow a precise sense of what is “significant” or “insignificant”. Such terms have previously been applied in an unscientific sense to describe the distribution of images.

3.1 Relational Theories

These are probably the most common hypotheses. “Relational theories” state that carvings can be “associated” with each other, giving them a new meaning. Locally, the differentiation of various series of carvings at the site area has been attributed to different cults: hypotheses have been made relating the concentration of specific themes to geographic elements like summits (cult of mountains) or lakes (cult of water). The whole site around Mont Bego summit is considered an open-space sanctuary (Lumley, Échassoux 2011). Despite this frequently claimed “sacredness”, the final conclusions of several authors are quite different (not to say opposite).

The only statistical work, investigating whether there is a structure to the images on the carved surfaces, was done by T. Serres (2001). Based on the study of 3,704 carved surfaces, representing most of the carved rocks known in “Les Merveilles” and “Fontanalba” sector, T. Serres observed repeated associations of types of carvings. According to him, there are 81 types of “significant associations”. These “significant associations” are interpreted as “ideograms” and, more recently, served as a basis or the proposition that these carvings formed a “proto-writing” (Lumley, Échassoux 2011).

3.2 Non-relational Theories

“Non-relational theories” state that the presence of a carving on a rock is independent
from any other carving: on a rock, carvings have no more meaning taken together than they have considered one by one. These assumptions are based on the weakness of the coefficients of correlation between carved themes, weak regularities in the organization of images on carved surfaces, weak tendency to avoidance between carvings on the carved surface and a distribution of carved rocks that is close to Poisson law.

In the “Les Merveilles” sector, for a sample of 90 carved rocks chosen, L. Barral and S. Simone (1990, 1991) calculate that coefficients of correlation between carved themes are lower than the significance threshold (ρ ≥ 0.21). Based on these results, L. Barral and S. Simone conclude that the presence of a theme on a carved rock is independent of the presence of another theme (which is a statement of Poisson’s law). Furthermore, the number of carved themes on carved rocks also follows the Poisson distribution: there is only one population of carved rocks.

In the “Fontanalba” sector, C. Chippindale (1988) observed weak regularities in the organization of images on the carved surface. His statement is in agreement with C. Bicknell’s (1913), for whom there were only a few examples of organized distributions of images. C. Chippindale also experimented with simulations to assess whether actual image distributions were random. He concludes that carvings are a little bit more organized in the real case than in the simulations but less so than in an organized composition.

For both, L. Barral and S. Simone (1990, 1991) on the one hand, and C. Chippindale (1988) on the other, the only strong regularity seen in carvings on a surface is in their orientation. Most – i.e., significantly higher than in a Normal distribution – are oriented toward the top of the carved surface.

For followers of “non-relational theories”, concentrations of carvings and carved themes are linked to the presence of suitable rocks for carving and areas of grazing (Chippindale). L. Barral and S. Simone conclude that carvings served to indicate areas of grazing for shepherds. According to C. Chippindale, the agro-pastoral hypothesis is also the most plausible.

3.3 Synthesis

As we see, these two points of view are opposed. On the one hand there’s a sacred and propitiatory hypothesis (T. Serres, H. de Lumley, A. Échassoux). On the other hand, there’s a secular and utilitarian hypothesis (L. Barral, S. Simone, C. Chippindale). “Non-relational” theories tend to consider the entire corpus of images and to, at least implicitly, use statistical approaches, considering the most common image types and the extent of variation. “Relational” theories, on the other hand, tend to focus on particular, often unique, images and to build an interpretation on that basis. In this sense, it confuses “exceptional” with “significant”. In both cases, distances, proximities, threshold of significance should be made explicit. Our work sheds light on which of these hypotheses is more credible.

4. Analysis

In this section we present some evidence that helps decide between the two hypotheses. Figures 2 and 3 show the external (geographic) and internal (type of rock, colour, etc) factors that help explain the concentrations of carvings or themes. To determine whether the differences (of location, accessibility, altitude, iconographic composition, etc.) between each carved rocks and carved themes are significant, we use Dunn’s post test: a non-parametric multiple comparison test based on pairwise comparisons. Its value is based on the differences between the sum of the ranks for each carved theme in comparison with the theoretical average (depending on the number of carved themes and the size of the samples). The result of Dunn’s test is a classification of carved
themes into different groups. In this part, we also consider the results of superimposition analysis.

4.1 Synchronic vs. Diachronic

In both “Relational” and “Non-relational” theories, diachronic information has been little-used to date. However, the study of superimpositions (Arcà 2009; Huet 2012), shows that successive periods are characterized by particular motifs and that there is a unidirectional trend (Fig. 4); weapons carvings (daggars and halberds) are among the most recent. The distribution of carvings on carved surfaces shows that a majority of carved rocks only received one carving. Moreover, regarding T. Serres work (2001), there’s a strong correlation between the total number of carvings on a rock and the total number of “associations” (R²=0.8). In other words, grouping of two or more carvings in an “association” could be accidental.

Multivariate analyses can show a seriation of carvings. In figure 5, a Correspondence Analysis, regarding images of yokes, for 254 carved rocks; it is possible to observe a «Guttman effect» (or arc effect). This characteristic form along the axis representing the first eigenvalues indicates a seriation of rocks from those with the most schematic carvings (characteristically

Figure 2. Extrinsic factors explaining concentrations of engravings or engraved themes (results of the Dunn’s test).

Figure 3. Intrinsic factors explaining concentrations of engravings or engraved themes (results of the Dunn’s test).

Figure 4. Examples of superimpositions of engravings in “Les Merveilles” sector (no scale).
of the Les Merveilles sector, on the left) to the most naturalistic forms (characteristically of the Fontanalba sector, on the right). According to the postulate that graphical elements appear and disappear gradually and linearly during time, seriation is equivalent to a process of evolution; the time-direction of this evolution is given by characters (naturalistic human figures) that hold halberds (Fig. 6). Considering the base of the blade of these halberds, more often convex than rectilinear, these halberds are more recent than the halberds which are represented independent of human figures (isolated halberds). As the same figures that hold halberds are associated with naturalistic yoked animals (bovids) it must be the case that these naturalistic animal images are contemporary with the later halberds and so the naturalistic animals are later than the schematic animals.

Considering now which themes are often represented on the same rock as another theme, we find weak positive correlations and weaker negative correlations (Fig. 7). This can be interpreted as a “palimpsest effect” meaning that surfaces were used continually across many periods. This “palimpsest effect” is particularly important on red rocks, rocks that seem to have been preferentially chosen during all periods for carvings (Fig. 8). But,
when we considered the coefficient of variation \((cv)\) of these correlations between themes, we found that fringed figures are well correlated with regular shapes and geometric figures (Fig. 9). Besides their good level of “association”, confirmed by the \(cv\), study of superimpositions tends to prove that fringed figures, and some regular shapes and geometric figures, belong to the same chronological period: probably the oldest one. Other confirmative elements can be found in spatial distribution.

4.2 Spatial Proximities to Geographic Elements

Early in the study of Mont Bego, concentrations of carvings and carved themes were identified (Bicknell 1913). One century later, the realization of a GIS database has allowed us to locate precisely the concentrations of carvings and carved themes and to establish which geographic factors might explain such concentrations. Two-thirds of the 4,274 carved rocks have been recorded with a precision better than 5 m in the \(x, y, z\) directions; one-third with 5 to 20 m precision.

The DEM (Digital Elevation Model) has been realized with a 10 m average precision. It allows cost-distance calculations for the whole site but is still insufficiently precise for systematic viewshed analyses from carved rocks. To calculate distances from carved rocks to various geographic elements (summits, pastoral paths, lakes, etc.), we constructed a simple isotropic cost-distance model based on slope and avoidance of water surfaces: it takes twice as much time to travel on a 12°-22° slope than it does on a flat or almost flat slope. Slopes equal to or steeper than 50° are considered impassable. Because our model excludes steep and water surfaces, some carved rocks, due to the imprecision of their location or immersed after dam construction, are not considered in the analysis. These rocks represent less than 3% of the whole corpus; the missing data can be considered negligible.

Walking speed has been considered to be 5 km per hour. This cost model has been compared with real-world movement and can be considered slightly optimistic (time travel slightly underestimated). The hypothesis test used to compare differences of spatial proximity between carved rocks, on which lie carved themes, and geographical elements is Dunn’s post test.

Figure 9. Distribution of the coefficient of variation (CV) for the coefficients of correlation of Matrix 1.

Figure 10. Location of engraved rocks with fringed Figures and the other engraved rocks in Les Merveilles sector.
Geostatistical analyzes show that fringed figures are significantly closer to permanent streams and lakes than other carved themes (Figs 10 and 11) and that there is no correlation between concentrations on carved rocks and summits, contrary at what “Relational” theories state.

However, in the case of summit, rather than cost-distance calculation, the study of visibility and topographic prominence (Llobera 2001) will be interesting to do, to confirm if proximity to summits is a variable able to explain carvings concentrations. Analyses also show that there is no correlation between grasslands or pastoral paths and concentrations of carvings/carved themes, contrary at what “Non-relational” theories state.

5. Conclusions

Quantitative methods and the use of GIS for important open-air rock art sites allows the highlighting geographical organizational principles (Alexander 2011, Huet 2012). These methods allow an easier access to data, systematic analysis and the standardization of results. However, in all these analyses, we only considered carved rocks. Recognition of suitable rocks which have not been decorated, and their recording in the same way as carved rocks, will be important to distinguish which factors are variables which are more able to explain the presence/absence and concentration of rock art. This could be done for a sample area. To investigate in-depth the spatial organization of the carvings, we must also continue to consider the whole recorded corpus; improving the DEM resolution to study visibility aspects and ensure that all analyses are routinely conducted across all potentially archaeologically-relevant subsets of the data.

Acknowledgements

I am grateful to Craig Alexander for his corrections and comments.

References


Abstract:
In the last decade integrated non-destructive survey methods have offered new possibilities for the identification and recording of buried archaeological sites, whilst reducing to the minimum destructive intervention and the cost of investigating large archaeological sites. Using the Roman site of Ammaia (Marvão, Portugal) as a case study, this paper discusses the ways in which the informative potential of archaeological datasets derived via multi-method surveys can be best exploited through the use of computational methodologies. The first part of the paper offers a brief account of how GIS-based data integration, data fusion, 3D reconstruction and visualisation have been used until now to enrich the interpretation of archaeological datasets collected at Ammaia. The second part discusses in more detail the yet largely unexplored potential of urban network analysis to enhance the interpretation of ancient street networks revealed with non-destructive survey methods, as well as some of the problems involved in the interpretive process.

Keywords:
Integrated Geospatial Approaches, Geophysical Survey, GIS, 3D Reconstruction, Data Fusion, Human Movement, Spatial Network Analysis, Space Syntax

1. Introduction

In the last decade advances in archaeological survey and remote sensing technologies (archaeological geophysics, aerial photography, satellite imagery, LiDAR etc.) have offered new possibilities for the recording, visualisation and analysis of buried archaeological sites. Nowadays, different geospatial techniques can contribute pieces of invaluable information on archaeological remains still laying underground in finer resolutions and with higher accuracy than ever before. In parallel with these technological developments there has been a growing awareness of the benefits of integrating the results of various archaeological survey and remote sensing methods to enhance and validate interpretations of buried archaeological sites. Such multi-method approaches have proved particularly successful in the context of urban archaeology, as sometimes they permit the identification of fairly complete layouts of urban and sub-urban areas offering an image of ancient towns that would have been impossible to obtain in the past merely by means of excavation.

Using the Roman site of Ammaia (Marvão, Portugal) as a case study, this paper aims to demonstrate that these essentially new datasets have a great informative potential that can be best exploited with the use of computational methodologies. Ammaia has been the main “laboratory” of experimentation in the context of “Radiography of the past” (Radio-Past -www.radiopast.eu), a multi-disciplinary and international2 EC funded project which aims to apply and further develop non-destructive survey methods to valorize complex archaeological sites (Van Roode et al. 2012).

2 Four academic institutions [the University of Évora (Portugal), Ghent University (Belgium), the University of Ljubljana (Slovenia) and the British School at Rome] and three companies (7Reasons Media Agency (Austria), Past2Present (The Netherlands) and Eastern Atlas (Germany)] participate in the project.
To date, archaeological research at the Roman site has offered an unusually large amount of diverse multidimensional datasets derived from geophysical survey (magnetometry, electrical resistance, GPR survey), topographic survey (total station, DGPS survey), aerial photography, excavation, terrestrial laser scanning, and 3D reconstruction of buried architectural remains. This very rich data collection makes Ammaia an illuminating case study for examining how archaeological data derived mainly by the application of non-destructive integrative geospatial approaches can be further interpreted, analysed and communicated to scholars and the public with the use of computational tools. The first part of this paper offers a brief background of the research methodology employed until now at Ammaia. It presents in summary how GIS-based data integration, data fusion, 3D reconstruction and visualisation have been used to interpret the datasets collected at the Roman site. A detailed description of these aspects of the project is beyond the scope of this paper, however, since they have been thoroughly discussed in a recently published edited volume (Corsi and Vermeulen 2012). The second part of this contribution looks into the yet largely unexplored potential of urban network analysis to enhance the interpretation of ancient street networks revealed with non-destructive survey methods, as well as some of the problems involved in the interpretive process.

2. A Brief Research Background: Data Collection, Integration and Visualisation at Ammaia

2.1 Data Collection

The Roman town of Ammaia is situated in the modern district of Portalegre (Portugal) in close proximity to the village of Marvão.
and the Spanish borders (Fig. 1). As Ammaia never hit the headlines of ancient writers, we have no record of the history of the town; its foundation has been placed in the Augustan age (between the end of the first century BC and the beginning of the first century AD) solely on the base of archaeological finds.

To date, only few parts of the ancient town have been brought to light by excavation. Until 2008, what was known of the site was an extensive part of the suggested main gate of the Roman settlement (the Southern gate: “Porta Sul”: Fig. 1, n. 3), a stretch of the town walls at the SE corner of the wall circuit (Fig. 1, n. 5), the podium of the Forum temple, and a few segments of the portico that surrounded the Forum and delimited the main market square (Fig. 1, n. 1). A small sector of the public thermal baths at the south of the Forum complex and some portions of a housing sector next to the eastern walls, located exactly below the building that nowadays hosts the local museum, were also unearthed in the 90s (Fig. 1, nn. 2, 4). In addition, excavations at the Baths have been resumed stratigrafically in recent years (campaigns 2008, 2009, 2011: Corsi 2012, 162-164).

As it is largely unexcavated, the site has been chosen by the research team of the Radio-Past project as the main “open-lab” for testing the application of multi-method non-destructive survey. Besides the study of existing aerial photography and the geoarchaeological survey which were carried out between 2001 and 2006, most work since 2008 has concentrated on geophysical survey. First, tests were performed with a Ground Penetrating Radar (GPR) in a small sector of the Forum and, given the good results achieved, 4.1 hectares of the estimated “intra-mural” area were surveyed with a fluxgate gradiometer in 2009. During the years 2010-2011 the total coverage of the lower part of the town, which is thought to have effectively been urbanised\(^3\), was accomplished with magnetometry (Johnson 2012), while in some selected areas three different geophysical surveys (high resolution GPR, magnetometry and earth resistance surveys) were performed. In addition, the modern national road, which cuts through the site, separating it in almost two equal halves, has been surveyed with GPR, and important segments of structures have been identified under the road pavement (Verdonck and Taelman 2012). Finally, in 2010 and 2011 magnetic mapping was performed in extensive areas outside the city walls using a wheeled array consisting of six Foerster fluxgate gradiometers; many archaeological features belonging to the suburban areas of the Roman city were registered during this survey, including remains of roads apparently leading to the city gates, building complexes and a necropolis (Meyer et al. 2012).

2.2 Data Integration

The integration of the geophysical, topographical and geoarchaeological datasets collected at Ammaia in a GIS environment was an important first step for the interpretation of the urban remains of the Roman town. Data integration in a GIS significantly facilitated the interpretation of geospatial information, either by simply enabling the combined visualization and interrelation of different datasets, or via the use of more sophisticated data fusion techniques. The comparative evaluation of geo-referenced results obtained with various geophysical methods (magnetometry, resistivity, GPR) which recorded different physical parameters of the subsoil provided a better definition of the location and geometry of archaeological bodies (cf. Piro et al. 2000; Nuzzo et al. 2009). On many occasions simple two-dimensional overlays of vector interpretations of magnetometry data upon GPR maps offered important complementary information (Fig. 2) or highlighted inconsistencies in the datasets examined. Furthermore, continuous data fusion techniques (cf. Kvamme 2006, Ogden et al. 2009) were also employed with various

---

\(^3\) The steep slope of the Malhadais hill is also considered to have been enclosed into the wall circuit for strategical reasons, although no built structures have been identified in this area.
levels of success depending on the nature and number of combined datasets as well as data quality issues. Such methods enhanced the visibility of geophysical anomalies especially where good quality high resolution datasets from magnetometry, earth resistance and GPR surveys were at hand, such as those obtained at the area of the Forum (Verhegge 2012; Paliou 2012). Finally, the integration of 3D vector interpretations of GPR data with topographic information and digital drawings of excavated remains gave a very comprehensive and informative picture of the nature, shape and position of the preserved visible and subterraneous architectural features4 (Fig. 3).

The integrated data derived by the multi-method geophysical survey at Ammaia pieced together a fairly detailed image of the subsurface of the intra mural area (Fig. 4) offering a general impression of the town plan. The latter has been conceived as an orthogonal grid, the idealized planning scheme being reconstructed with three rows of rectangular insulae with sides in a proportion of 2:1 or c. 90:45m (2 ½ : 1 ¼ Roman actus) on the northern and southern sectors of the town, and two central rows of blocks, a bit wider along the short side (with a proportion of 2 ½ : 1 ½ actus: Corsi et al. 2012). On the geophysical map the insulae of the town are clearly delineated in most cases and public (e.g. a basilica, a sanctuary) and private (e.g. houses, shops) buildings can frequently be identified. Details of the interior of built units, such as rooms, entrance locations, and peristylia, can be distinguished on some occasions, especially when high resolution GPR data5 are also available (Verdonck and Taelman 2012).

4 ArcGIS 10 was used in this case because it can incorporate certain types of 3D geometry, as well as import models created with 3D modeling packages (e.g. CAD).

5 In the case of high resolution GPR survey a transect spacing of 5cm was used (Verdonck and Taelman 2012).
2.3 3D Reconstruction and Visualisation

GIS data integration and fusion has, thus, greatly facilitated a traditional analysis of geophysical datasets derived from complex urban sites, where emphasis in interpretation is placed on the identification of public and private buildings, fortifications, gateways, and street segments (cf. Benech 2009, 89). Besides GIS based analysis, however, another computational approach that has extensively been exploited to enhance data interpretation in the context of Radio-past project is the 3D digital reconstruction and visualization of visible and subsurface building remains. Two are the main aims of this process: First, to attempt a
three-dimensional interpretation of the Roman town that incorporates and combines evidence from available geophysical, geoarchaeological, topographical and excavation datasets with comparative archaeological information derived from contemporary Roman urban sites. Second, to communicate in a comprehensible way the results of the project to a non-specialist audience, so as to encourage public understanding and engagement with the hidden archaeological environment of Ammaia.

The reconstruction of the building remains of the Roman town has been attempted at various spatial scales. For some areas of the site, such as the Southern Gate and the Forum, reconstruction has been based on walls that are still visible in situ, geophysical survey results and comparative architectural data from other Roman sites of Lusitania (e.g. Conimbriga), but also on stratigraphic information, obtained via excavations and focused ground truthing operations, including small trenching and augering (Corsi et al. forthcoming). In this way the process of digital reconstruction encouraged a re-evaluation and re-interpretation of all available evidence, elaborating many of the structural details of the buildings under study. In addition, a much larger scale visualisation of the Roman town has been attempted (Fig. 5) that seeks to give a broader view of the site in its cultural and natural landscape (Radio-Past Team 2013). In this case 3D modelling is supported mostly by information found in the geophysical maps, while architectural local features and decorations, known from better preserved and more extensively excavated archaeological sites of the region, are used to complement a tentative picture of the provincial Roman town. On all occasions, decisions made in the course of the reconstruction process are documented on the online blog of the project, and are discussed, commented and contested by the project members to ensure transparency and a scholarly approach to the problems associated with the proposed data interpretations (Klein et al. 2012).

3. The Application of Urban Network Analysis to the Roman Site of Ammaia

3.1 Analytic Approaches to Human Movement in Archaeological Built Environments

The potential of urban network analysis to offer an additional dimension to the interpretation of geophysical datasets is an important research question that is currently being explored in the framework of Radio-past project. Integrated geospatial surveys of ancient towns provide an abundance of data that not only elucidate the form of past urban environments, but they could also be amenable to a quantitative examination by means of urban network analysis. Such an approach could offer insights into social aspects of human movement and interaction in past built environments. To date, formal analysis of human movement has been very popular in archaeology, especially in the realm of landscape studies. GIS-based least-cost path analysis has long been utilised for
exploring socio-symbolic aspects of the natural environment that past people used to inhabit, looking into the ways in which human mobility may have affected choices on site location, and/or human communication and interaction at larger spatial scales. At the level of urban settlements analytic approaches to movement, for example Space Syntax techniques, GIS transportation network analysis or agent-based modelling, are less commonly employed, partly due to the difficulty in acquiring sufficient data on the form of ancient street networks. Such methods are usually concerned with identifying public urban areas that would have given increased opportunities for social encounters and, hence, were of potentially high social significance, e.g. the main thoroughfares in the street network, the passages that a visitor to the town was more likely to traverse, or the open public spaces the town dwellers would have more frequently used in the course of daily life.

Up until now, axial analysis, namely the graph-based technique that was introduced by Hillier and Hanson (1984) to explore the ways in which the spatial configuration of a street network can facilitate or discourage human movement, co-presence and interaction, has met relatively few archaeological applications (see for example Ferguson 1996; Potter 1998; Kaiser 2000; Robb 2007; Craane 2007; Kaiser 2011). Axial analysis focuses on the topological properties of an urban network aiming at identifying the most “accessible”, and therefore most widely used street segments (Hillier and Hanson 1984, 82-142). The term “axial” refers to the first stage of the analysis in which the built environment is represented in terms of axial lines. These are the longest and fewest lines of sight that traverse each outdoor space in a continuous urban plan. At the second phase of the analysis axial lines are represented as nodes, and their intersections as links in a graph, which can be quantitatively described with graph-theoretic measures (Hillier and Hanson 1984, 82-142; Conroy 2001, 12). A main obstacle in the application of axial analysis in archaeology is that the urban networks under study should be continuous and complete (all spaces must be described in terms of axial lines) which is hardly ever the case for the great majority of ancient urban spaces. The same limitation applies to other formal analyses of human movement in a street network used by archaeologists, such as GIS-based network analysis and agent-based modelling. GIS-based network analysis, besides topological information, can incorporate a great amount of contextual information related to socio-economic data and the use of space, taking into account also factors such as the direction, topography and cost (e.g. distance, time, energy) of movement (Fischer 2003; Branting 2004; Conolly and Lake 2006, 236-238). On the other hand, agent-based models (ABM) aim to study human mobility at medium and small spatial scales from the bottom-up, by examining the non-linear formation of collective patterns of movement that emerge as a result of interactions among individuals. Such models could comprise of agents with advanced cognitive abilities (e.g. memory, knowledge of the environment, clearly defined aims) and a well-defined course-determining mechanism involving goals, learned paths and destinations (e.g. Haklay et al. 2001; Kurose et al. 2001). ABM of this kind follow a set of rules that are programmed by the user and should be distinguished from agent-based approaches inspired by Space Syntax6 where the behaviour and actions of agents are driven solely by the configurational properties of space (Turner and Penn 2002). Despite the fact that GIS spatial network analysis and complex ABM can allow for numerous factors that influence the behavior of mobile individuals, they have been rarely applied to archaeological built environments (Branting 2004; Altaweel and Wu 2010). This is possibly due to their computational sophistication and their increased requirements for socio-economic data that are hard to obtain for historical and prehistoric societies.

That said, recent technological advances in archaeological geophysics and remote sensing presently provide more opportunities

---

6 Such analyses can be implemented with UCL Depthmap.
for obtaining comprehensive information on urban layouts at medium and large spatial scales, permitting and facilitating the application of quantitative techniques that can describe and identify key elements of ancient street networks, such as major through routes and patterns in the use of space (cf. Kaiser 2011). It is noteworthy, however, that archaeological research to date has taken very little advantage of these possibilities with only few exceptions; Branting’s (2004) study on human movement and transportation at the Iron Age site of Kerkenes Dag is one of the rare – if not the only – application of GIS-T (GIS transportation network analysis) to a prehistoric urban network that was identified mainly by extensive geophysical surveys. More recently, Kaiser (2011) calculated some basic network measures of the street grid of the Roman part of the ancient site of Ampurias (Empúries), which has been revealed to its greater extent through aerial photography (Kaiser 2011, 194), and Stöger (2011) considered the results of geophysical surveys in her axial analysis of Ostia. Furthermore, Benech (2007) and Morrow (2009) discussed the use of Space Syntax techniques for the analysis of geophysical datasets, nonetheless their studies focused on the scale of buildings, rather than on the settlement level.

3.2 Axial Analysis of the Street Network of Ammaia

The integrated geophysical survey at Ammaia provides then another opportunity to examine how geophysical datasets collected at a complex Roman urban site could be further explored and interpreted with a formal analysis of the street grid. As mentioned above, the urban layout of the Roman town has been made known mainly by the results of geomagnetic prospection, which, however, “flattens” different phases of development and combines walls and features that potentially belong to different periods into a single image (Corsi et al. 2012). Transformations over time in the street grid, which was probably owned by the civic government according to Roman laws (Kaiser 2011, 21), should have been less common, however, than changes to the configuration of privately owned space in the town’s insulae; roman legal codes also required that property owners respected the spatial limits of the urban grid making sure that the portion of the street that traversed their houses was passable (Kaiser 2011, 23). Nonetheless, the image of the street network we get from the geomagnetic survey is not always clear; in some instances features appear to interrupt, cross, or lay under/over the linear continuations of the road surfaces of an idealized street grid. On many occasions such features are linear and are placed along the sides of the streets without completely blocking the areas that can be attributed to street segments in the geophysical map. It is possible that these features underlay or worked in conjunction with the streets like sewers, sidewalks, and public fountains7 (Corsi et al. 2012). In two or three occasions south and east of the Forum area, however, it seems possible that built elements could have significantly narrowed or completely blocked public passages. Since no pre-Roman settlement evidence exists in Ammaia, it is possible that either these streets were encroached upon in a later phase by buildings8 or that no street existed there even in previous phases of occupation e.g. there was an intended deviation from an idealized plan ever since the street grid was laid (Corsi et al. 2012). Perhaps in the future excavation or GPR survey could further enlighten this issue by offering more information on the depth and dating of the structures in question. At the moment, since the occasions where streets may have been completely blocked by built features seem to have been few, the use of alternative reconstructions of the street network and the application of spatial analysis to each different

7 Geophysical prospections in Italica have shown the presence of several public fountains located near the corners of some insulae on the pavement of the central street of the city (Keay and Rodríguez Hidalgo 2010).
8 This is a phenomenon which has been proven by excavation in many Roman towns of the Mediterranean, including recently in the riverside quarters of nearby Mérida where from the 5th century onwards the town insulae start to be spoliated and compartmented (Alba Calzado 2004, 216).
case could elaborate at least the extent to which such uncertainties in the data could affect the proposed interpretations. In the analysis that follows we assume that the street network discerned in the geophysical map belongs to a later phase of the town, if it does not correspond also to the original street grid.

With the above limitations in mind, a first interpretation of the street network of Ammaia with Space Syntax techniques has been attempted using Depthmap, a piece of software developed at University College London. The software can calculate a variety of graph measures that have been used in the framework of Space Syntax theory: the “connectivity” (degree centrality), which indicates the number of intersections of each street to all other streets, the “choice” (closely comparable to betweenness centrality), which aims to determine whether a street occurs frequently in shortest path routes calculated for all pairs of possible origin-destination locations (Turner 2007), and the “mean depth” that expresses the sum of the topological distance of a street to all other streets in the network divided by the number of streets minus 1. Furthermore, graph measures that have been introduced by Hillier and Hanson (1984), such as the “integration”, can also been calculated using DepthMap. The “integration” is a normalized measure of the mean depth. Street segments that correspond to axial lines which are of a small average distance to other lines in the spatial configuration are termed “integrated”, or, if the opposite is the case, “segregated”. Integrated and segregated street segments are considered indicative of the distribution of traffic volumes in the network with the former signifying the most frequently used passageways. To date it has been demonstrated that in modern built environments there is a good correlation between integration indices and pedestrian movement counts (Hillier et al. 1993; Conroy 2001, 12; Turner 2003). Most successful predictions of traffic volumes have been derived when only lines that are three topological steps away (“radius three integration”) have been considered in the analysis (Conroy 2001, 12). Nonetheless, in this case it has to be recognised that correlation indicates association and not necessarily causality, and in this sense it cannot be precluded that patterns of mobility have been shaped under the influence of factors other than spatial configuration, which are not taken into account in Space Syntax approaches. The Space Syntax centrality measures described above (“integration”, “connectivity”, “choice” and “mean depth”), could potentially reveal the major thoroughfares and the most widely used street segments in ancient street networks as applies to modern built environments. Furthermore, the examination of these measures together with information on the various functions and human activities associated with private and public buildings can suggest some of the society’s ideas about the proper use of space and the social processes that created the spatial layout of ancient cities (cf. Kaiser 2011).

In the case of Ammaia the calculation of all centrality measures shows that the NW-SE streets have high degree of centrality when compared to the SW-NE streets (Fig. 6, 7, 8, 9). This would suggest that NW-SE streets were important for channeling movement and transportation through the Roman town (cf. Kaiser 2011, 49). The integration of the results of network analysis with contour data offers a plausible explanation of why this pattern has emerged (Fig. 10): it appears that the city grid was laid in such a way so that the NW-SE streets are parallel to the topographic contours; such arrangement would have greatly facilitated pedestrian and vehicular movement towards NW-SE direction. On the other hand, movement along SW-NE streets that are laid perpendicularly to the contours would have been less easy due to more abrupt changes in slope. Furthermore, the calculation of topological distance from key areas of the street network, such as the forum or the city gates (cf. Kaiser 2011, 52-55) could indicate the

---

9 This is more apparent at the NW where there is less evidence for human intervention (e.g. terracing) in the landscape after the Roman era.
Figure 6. The “Connectivity” of the street grid of Ammaia. Frequently used streets are indicated with high values.

Figure 7. “Choice” measures for the street network of Ammaia. Values suggest how many times each street is encountered in shortest path routes calculated for all pairs of possible origin-destination locations (streets).
Figure 8. The “Mean Depth” of the streets of Ammaia. Low values indicate short topological distance and a high level of integration.

Figure 9. “Integration HH (Hillier-Hanson)” for the streets of Ammaia.
Figure 10. Connectivity measures for the street grid of Ammaia integrated with 1m. topographic contours of the modern terrain.

Figure 11. The “Step Depth” from Porta Sul. Low values indicate high topological proximity.
streets that the visitor of the town would more likely traverse. Figure 11 shows the “Step depth” from Porta Sul, namely the topological distance between the main gate of the town and each street in the network, suggesting perhaps the routes a traveler coming from the provincial capital of Emerita Augusta (Merida) would encounter.

More insights into the social life and urban organization of the Roman town will be precipitated, if graph measures are associated with the interpretation of buildings (e.g. public and private) identified in geophysical maps, a process that is still in progress. Furthermore, other approaches to network analysis, including a GIS approach that takes into account metric factors that influence movement, could also be incorporated in interpretations, and combined or contrasted with the results of Space Syntax approaches. Although the analysis of human movement through the open public spaces of Ammaia could be complemented with alternative analytical methods, the case study presented above already suggests that urban network analysis could be successfully applied to ancient street networks revealed with integrated non-destructive survey methods and that such approach has the potential to enrich our interpretations of ancient urban sites.

4. Summary and Conclusions

This paper discussed the ways in which computational integration, visualization and analysis of geospatial datasets collected at a complex urban settlement can augment archaeological interpretations. The greater advantage of the multi-method research methodology discussed above is that it assembles fragmented data into a more informative whole, enhancing the interpretive value of individual datasets, whilst reducing reducing to the minimum destructive intervention and the cost of investigating large archaeological sites. Geospatial data integration in a GIS environment is nowadays a routine procedure in archaeological practice that over the years has become more effective and sophisticated owing to the more efficient processing of prospection data and the improved analytical and 3D visualisation capabilities of GIS software. On the other hand, the benefits of large scale 3D reconstruction and modelling of both visible and subterraneous building remains of ancient settlements have started to be explored only in recent years. The 3D visualization of Ammaia, taking advantage of the results of extensive geophysical surveys, as well as geoarchaeological, topographic, excavation and comparative archaeological evidence, has enhanced data interpretations and encouraged further communication and scientific debate among the project collaborators. At the same time it has proven particularly useful for presenting research results and the interpretations proposed by the scientific team of the project to the public. Finally, the interpretive potential of the application of urban network analysis to a street grid revealed mainly with integrated geophysical surveys has also been considered. The use of axial analysis to the street network of Ammaia gave some useful insights into human mobility through the streets of the Roman town and should be seen as a starting point for future work. Similar analyses have been applied also to largely excavated Roman towns (Kaiser 2011) making possible the quantitative description and comparison of distinctly different examples of Roman town planning.

The further examination of Roman sites with integrated geophysical surveys promises to offer suitable datasets that could permit more formal comparisons of this kind in the future. It is noteworthy that in our study the large scale 3D visualisation and network analysis of geophysical datasets were greatly facilitated by the fact that the urban layout under study clearly follows a regular grid and was the result of careful planning. In cases of unplanned street networks the identification of passageways using the results of geophysical prospection presents...
more difficulties, as there is greater uncertainty regarding the form and chronology of the street grid. Nonetheless, as applies also in the case of Ammaia, a combination of geophysical surveys with targeted excavation and ground truthing could resolve some of these problems and offer a more comprehensive image of buried ancient townscapes.

Acknowledgements

The authors would like to thank all the collaborators of the Ammaia project and especially Frank Vermeulen, Paul Johnson, Michael Klein, Cornelius Meyer, Sara Persichini, Devi Taelmann, Nicola Schiavottiello, Lieven Verdonck, Jeroen Verhegge, and Guenther Weinlinger. The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement nº 230679, under the action Marie Curie – People IAPP, with the Project entitled ‘Radiography of the past. Integrated non-destructive approaches to understand and valorise complex archaeological sites’.

References


Fischer, M. 2003. “GIS and network analysis.” In


1. Introduction

Evidence of spatial discontinuity is an important source of knowledge for settlement transformation issues. Change in the topographical location of inhabited spaces could often reveal structural alterations in society. Stabilization of settlement structure in High Middle Ages in Bohemia is a typical process where the spatial and temporal continuity or discontinuity reflects important changes in the organization of society.

Written sources and surface artefact surveys offer a basic model for the high medieval settlement transformation process. The less stable and more dispersed early medieval hamlets are superseded by nucleated villages which often persist in their original location till today (Klápště 1991). Although there are indications of a more complex process documented at several excavated sites, we need tools which could detect more gradual changes in topography.

Artefacts obtained during surface field survey are commonly used for the recognition of on-site or off-site topography. Using GIS and spatial statistics could not only increase understanding of the static spatial structure but also reveal the diachronic development.

2. Methodical Framework

The first question is how the present (observed) topography of particular artifacts reflects past settlement activities.

The typical recorded feature - a large concentration of ceramic sherds within small area - could be linked to the former habitation units (houses) or waste pits or other sunken features usually found close to the houses (Foard 1978; Schoefield 1989).

Second methodical question concerns manuring practice which could be assumed for all periods under study but largely in a different extent (Fentress 2000; Hejcman – Vilém 2010). We could assess later fertilizing activities by 19th century potsherds and cadastral maps from the years 1826-1843. Here we can prove that the spatial extent of early modern brown earthenware fragments on the fields could not be traced to any contemporary settlement.
The recognition of a low intensity habitation site from a high intensity manuring area is not possible for early medieval period only by the means of field surface survey. For the purpose of this paper we can probably assume that in our study area the fertilizing in earlier periods was focused only to the intensively managed fields in a short distance from the houses.

The most important change in the manuring practice probably took place in Bohemia sometime between the 13th - 14th centuries with some regional differences. The close spatial relation between the habitation area and intensively managed arable fields is abandoned due to colonization and adoption of new and until then unused areas for fields. The chronology of this change is very obscure but could be roughly traced due to the fact that for the same period we observe a change in the use of raw material and an elaboration of the ceramic fabric (Klápště 2005, 170-208; Žemlička 2002, 63-76, 615-625).

Third factor to be taken into the consideration is taphonomy. Later environmental or anthropological processes on sites could transform the spatial structure of artefacts on the surface of the fields and obscure their relation to the original activities. For example erosion could create unexpected accumulations in lower parts of hill slopes or reduce quantities in their upper parts.

All above mentioned activities are detectable via spatial properties of particular potsherds alone or grouped into larger survey units. The key feature of the presented method is quantifying the dispersion of pottery fragments within detected chronological categories. Obtained values are then used for the interpretation of activities which left behind the artefact collections.

3. Sites and Periods

As it was mentioned above, a great opportunity to detect and analyze changes in settlement is offered by the high medieval settlement transformation process. Surface artifact survey was undertaken on seven sites where early medieval settlement phase was known from earlier surveys or could be expected due to the presence of solely standing church of medieval origin (Fig. 1).

For all of them exists evidence for the presence of deserted medieval habitation activities but with only simple or no clearer picture of the settlement extent or its structure. All five studied churches have their early medieval or medieval origin verified (solely or in combination) by the existence of Gothic components, archaeological finds nearby or typically by written sources. Selected sites are located in different landscape types so it was supposed that also the structure of surface collections will differ.

We could suppose that these buildings were not built in isolation from any settlement entity and hence their current situation is a consequence of settlement transition. The church position is usually much more stable than that of settlement (Morris 1989).

The field part of the survey was realized in two ways with different intensity. The first one is based on fieldwalking with continuous GPS recording of the surveyed tracks. Only medieval or “probably medieval” finds are
collected with the position recorded. The intensity of the survey is usually not exactly the same everywhere within the studied area but could be easily quantified for example via the interpolation of survey lines within GIS. The second method is more intensive. The prospection is realized within squares of 50 x 50 meters and the position of the surveyor is traced by the GPS display where are visible the imported square polygons. No demarcation of survey units in the field is needed as well as in the case of previous method. Also this method is possible with one surveyor and the walked line is ca 330 meters on average in one square. Survey time is 10 minutes per square or shorter when the area is limited by the position at the edge of survey area - i.e. the field.

The main link to the deserted settlements is represented by the pottery or its spatial distribution respectively. Ceramic categories on studied localities were based mainly on the material characteristics which changed significantly through time and also due to the lack of the rims or other chronologically sensitive features. The collected pottery material was divided into four main and several additive groups. The oldest category was labeled K3 (early medieval, mainly 11th - 13th century AD). The second category - K4 (high medieval, 14th - 16th century) - represents an important change in the distribution of sherds which reflects the introduction of “distant manuring” on a large scale.

Modern ceramic classes were also collected. Although their manuring origin is almost sure it helps to detect the eventual taphonomical disturbances on the site. On all localities is the spatial distribution of early modern or later categories similar to their modelled random distribution. The third main group which is the easiest to distinguish by a glaze - K5 - consists of all detected pottery types of the early modern era. K6 (or subcategory K611) represents clearly distinguishable brown earthenware pottery manufactured after 1850.

4. Technical Framework and Application

Geographic information systems (GIS) were employed as a tool which significantly improves the possibilities of surface artifact surveys. Primarily it was used together with GPS for the demarcation of survey units and storing the collected archaeological and environmental data. Of particular importance is the correlation between location of finds and their attributes defining quantity, chronology and so on which is used for calculating the spatial structure of ceramic categories.

In this paper we present three methods which analyze spatial characteristics of potsherd distribution for the purpose of reconstructing their origin. This means to uncover the activity which left behind the ceramic scatters on the surface of the field.

Key feature of assessing the concentration (or dispersion) is the random distribution or in other words amount of randomness present. In GIS we can also model how the structure of particular data could look if they were distributed by chance (Fig. 2).

Three different methods which are presented in the subsequent part of the text represent three different approaches on the scale from exploratory visualization to statistical quantifying. The methods are Kernel Density, Standard Distance and Spatial autocorrelation (ESRI 2012).
**Figure 3.** Kozojedy site - output of the Kernel Density tool (Spatial Analyst tools / Density), search radius = 100. K3 - early medieval, K4 - high medieval, K5 - early modern, K6 after 1850.

**Figure 4.** Kozojedy site - output of Standard Distance tool (Spatial Statistics Tools/ Measuring Geographic Distributions) and spatial autocorrelation - Cluster and Outlier Analysis (Anselin Local Moran’s I - Spatial statistics tools/ Mapping clusters).
Application of Kernel Density represents an exploratory method very useful for the first view on the data distribution (Fig. 3). The first reason lies in the fact that this tool is a kind of interpolation. This could span the imperfections in the survey and visualize main trends in the data. Another reason concerns detection of taphonomy. In the case that the main concentrations of more or all categories are located in lower part of slopes we should take the increased influence of erosion into account.

For the purpose of using the Kernel Density tool the “collecting” squares were converted into the points (centroids). These points were then used for interpolation. The weight (physical) of all sherds of particular category collected within the square was used as the weighting factor in the computation.

When created for each category, it enables us to see the main spatial distribution characteristics. The modern and early modern ceramic classes have usually more dispersed pattern than the older. We also could assess the location of past settlement sites - in the case of classes in primary position. This is the main task for the subsequent methods - Standard Distance and Spatial autocorrelation.

Method of Standard Distance (SD) creates circles with radius equal to standard distance value. This value is computed as standard deviation of feature coordinates weighted (in statistical meaning of word) by specific feature attribute. This was in our case weight of pottery sherds of particular category in a survey unit. The circle is centered on the weighted mean centre of the feature set.

It is obvious that position and size of SD circle is affected by spatial distribution of the dataset. The clustered data is represented as small circles placed near high values. The circles for dispersed or even random data are much bigger and are placed close to the mean centre (Fig. 4).

This method is easy to implement and offers quick assessment of the data. A problem occurs when the dataset includes two or more clusters at different locations. Then the SD circle is located between clusters and is often bigger than random circle.

Third method - Spatial autocorrelation - could be used in two ways. As a global statistic it gives one characteristic for the whole dataset similarly to the SD method. On the other hand, it works with more sophisticated statistic background which could be described by the term “spatial autocorrelation”. This tool measures how a value in one location influences the values in the neighborhood.

In case of this paper, the Moran’s statistical test was used. The size of autocorrelation is represented by Moran’s Index at specific level of significance. The positive autocorrelation means that high values are located near other high values and the index is near to value 1. In case of random data the index is near to value 0. The threshold distance parameter influences the size of neighborhood of each point used in the calculation. The variation of this parameter helps to study the relationship between autocorrelation and distance. This relation can be shown in a graph called correlogram.

In case of clustered data the Moran’s Index is high at small distance and decreases with increasing distance. In case of random data, the Index is near zero regardless of distance. As shown in Figure 5, the comparison of correlograms is the main output for spatial distribution assessment.

It is necessary to stress that we should take into consideration the whole course of autocorrelation curves and combine them rather than rely only on one of them. They represent the trends which should be interpreted in context of the whole site.

In contrast with global methods, the local methods focus on significant changes in spatial
Figure 5. Correlograms - output of the Spatial Autocorrelation (Morans I) tool (Spatial Analyst Tools /Analyzing Patterns). Jenišovice site has slightly different scale on main axis for better visibility.

Figure 6. Relation between main scatters of K3 pottery and the position of church.
pattern. The so called hot and cold spots are identified.

In case of this research the hot spots need to be located. The local Moran’s I statistic was used for this purpose. The term hot spot means that the high values are accumulated at one location. The clustered data contains one or more hot spots while the random data does not contain any hot spot in the data set.

Detected hot spots were drawn to the map with weighted SD circles and the position of churches. Based on this visualization, the clustered and random data distribution could be distinguished (Fig. 4).

The outputs from Standard Distance and Spatial autocorrelation together created the base for decision about spatial distribution of dataset. As shown in figure 5, in the site Kozojedy, the correlogram shows cluster distribution for K3 and random for others. This result is confirmed by figure 4, where K3 has distinctly smaller circle of SD than random and there are hotspots found near the church. The other datasets K4, K5 and K611 have similar size of random and SD circle and there are no significant correlation between hotspots and church position.

To create outputs described above, the scripts were written which automatize the computation process. The Python scripting language and ArcGIS geoprocessing libraries were used. The graphs were generated using own Matlab scripts. The procedures were applied on 6 sites with 6 categories.

Main aim of the above-described methods is to discover the origin of pottery scatters of particular categories. If we know how the later manuring categories “behave”, we probably can in accordance with older surveys match the K3 categories on sites with places of former households or the waste areas in their proximity. The fact that the most concentrated and the most autocorrelated classes are probably in the primary position could also help to predict sunken archaeological features on the site which supply the topsoil with sherds. On the other hand we could, in the case of categories of unclear origin, support their interpretation by their dispersion characteristics.

5. Discussion

Typical problem of settlement behavior interpretations is the way how to validate them without excavation. One possibility of verification is to compare observed spatial distributions with the position of other monuments from the same period. Due to the focus of our survey on the hinterland of (former) medieval churches we can confront their situation and the location of main peaks of K3 pottery appearance. On all surveyed church sites proximity between them is clearly visible (Fig. 6).

Another procedure which was not used so far and probably could offer additional arguments for interpretation is the selection of fragments with sharp breaks. These usually tend to be very close to the sunken archaeological objects or layers.

Further research also has to deal with something what could be called a topographical threshold. This is the need to distinguish when the higher amount of dispersion is caused by rather disperse nature of settlement form and when this is an effect of “diffusive activity” - typically manuring.

The first situation could be represented by several smaller settlement foci in the survey area which are themselves defined by relatively dense scatters of ceramics. The second one could rise as a bigger area “normally” covered with potsherds. The problem is that the Standard Distance values (and often the autocorrelation curves) of the former tend to be fairly lower and then indicate rather the non-settlement activity.
For this and similar reasons it is not possible to highlight one particular method to be more efficient than others. Our work is so far limited to only several sites from somewhat historically and geographically different contexts. This fact is reflected in internal heterogeneity in the dataset on one hand and in the necessity to carefully interpret and compare results of all methods for each location. None from the above described procedures provides the same useful results on every site.

6. Conclusions

The main aim of this paper is to draw attention to several techniques which could offer further enhancement of information gained from surface artefact survey collections. The key idea is that there are two main properties concerning chronology. The first one is commonly used directly, when chronology is assigned according to the morphology, processing, fabric, etc. Dating could also be based upon spatial distribution which represents the second way.

We have suggested three methods which have different as well as similar aspects. Their difference lay in the statistical importance of the results. Spatial autocorrelation on its own enables assessment of statistical validity. Similar is the key feature when the observed distribution of each ceramic class is compared to its modeled random distribution. In our opinion, illustrative are also the differences in interpretation. Structures visualized by the Kernel density method are far more subjective when interpreted then that of spatial autocorrelation. Other reason supporting this statement is that the setting of bands in which the Kernel density map is calculated is subjective. Nevertheless, the meaningful value is distinguishable by trying several bands. The optimal value is different on each site according to its size and this again highlights the need of individual interpretation of the results on each site.

On the other hand there is a possibility to compare surveyed sites with each other according to the results gained by the referred methods. This is also the way how the major taphonomical distortion could be detected. When all Standard Distance values on the site are similar (or even high) or all autocorrelation curves are similar as the curve for random distribution, it is probably that the site underwent very intensive ploughing, topsoil shifting or similar destructive process.

Acknowledgements

Many thanks to Karel Nováček for the data from Kladruby site and to Regina Janíková.

References


Ecological and Social Space in the High Mountains in South Norway 8500 – 2000 BP

Espen Uleberg
University of Oslo, Norway

Ellen Anne Pedersen
Buskerud County, Norway

Abstract:
The first sites in the Norwegian high mountains appear shortly after the glaciers had melted; in the Hardangervidda around 8500 BP. A pine forest was established as high as 1250 m a.s.l. around 8700-8500 BP. The existence of a birch forest belt above the pine forest is debated. This is important when debating the sites in connection with hunt for only reindeer or also other big game. There are indications of shieling in the high mountains as early as the Late Neolithic. This subsistence change can be reflected in the allocation of sites as the Late Neolithic site distribution differs from the Mesolithic and Early Neolithic sites. This indicates that the perception of the landscape and the created ecological and social space had changed.

Keywords:
Stone Age, Norway, High Mountains

1. Introduction

The high mountains in Norway have traditionally been seen as a last refuge for the reindeer hunters of the North-European plain. Viewing the mountains today invites a feeling of unchangeability, and the annual reindeer hunt in the autumn can create a feeling of continuity from the first hunters in the area more than 8000 years ago to the modern reindeer hunters of today. This paper will discuss whether the site distribution in Hardangervidda reflects a change in the construction of ecological and social space on the basis of the ecological factors forest-limit and proximity to lakes.

Hardangervidda in South Norway (Fig. 1) is the largest peneplain in Europe. The area is around 6500 square kilometres. The average height is 1100 m asl., and the highest point of the plain is at 1721 m asl. The western part is dominated by rocky terrain that falls steeply down to the fjords on the west coast. The east is more flat with more vegetation and the landscape more gently descending into the eastern valleys. One of Norway’s largest glaciers, Hardangerjøkelen (1863 m asl.), is found in the Northwestern part, and today the whole peneplain is above the tree-line. It is a good habitat for reindeer, and one of the largest still existing wild reindeer herds in Europe is found here.

2. Prehistory in the High Mountains

The knowledge of the prehistory of the high mountain regions in South Norway is based mainly on surveys in connection with dam constructions for hydro electrical power, but also national plans for environmental protection of river basins (Indrelid 2009). An exception from this is the Hardangervidda Project, an interdisciplinary cultural research project (Johansen 1978).

The first sites in Hardangervidda appeared shortly after the glaciers melted around 8500 BP. The use of stone tools in this area, both flint and local quartz/quartzite, continued through the Bronze Age and into the Iron Age 2500/2000 BP (Indrelid 1994). Since stone continued to be a primary material for
tools, this period has been coined the *Stone Using Period* or the *Long Stone Age* (Pedersen 2011). Many of the sites are small with only expedient tools; no stratigraphy, no charcoal and no type specific artefacts. At other sites, typology and radiocarbon dating indicate certain time intervals for events at the site but not necessarily documenting the whole period the site was used. Even those that can only be dated within the Long Stone Age contribute to the understanding of the pattern of movement, activities and the construction of social space.

3. **Ecological and Social Space**

Ecological and social space is inhabited and created by the people in it. People with different economies will look for and perceive different things in their surroundings (Meløe 1989). A hunter gatherer looks for and perceives other elements than a pastoralist. The hunter will look for the movement of the game, the good places for hunting. The pastoralist looks for the best grazing areas, and how to contain his animals and guard them against predators. The two groups will have a different relation to the landscape, and we might infer their perception of the landscape through our spatial analysis of the site distribution. The ecological space is in this way dependent on the subsistence pattern.

The lakes in Hardangervidda are a dominating feature in the landscape. The lakeshore is a continuous space that has attracted events several times. Events separated in time but overlapping in space create a palimpsest where newer events disturb and overlap the older. What we can register is the debris, the left-over from the events. Agglomerations of such events registered on maps are the archaeological site distribution. The boundaries of each site are in many cases random or based on an educated guess. Activities that were once closely connected in time may be far enough apart in space to be grouped as two separate sites. At sites without stratigraphy the original distribution over the time axis has collapsed.

The artefact scatter may be the result of a short intensive period or several visits to the same place with longer intervals between them. C14-dates and dateable artefacts can give some points or intervals in time to which the events can be connected. Two diverging C14-results from the same site could indicate continuity, renewed annual visits to the same place, or two distinctly separate events. The difference between small and large sites is the number and diversity of events that took place there. What is left is artefacts and ecofacts, remnants of events distributed in space and now grouped as sites (Uleberg 2004).

The ecological space in the high mountain is constructed on the background of three main factors –the raw materials for tools, the forest limit and the big game.

3.1 **The Raw Materials**

The raw materials indicate movements both regionally and locally. 99% of the raw material at the Hardangervidda sites is flint, quartz/quartzite and slate (Indrelid 1994, 171). Flint shows contact with the coastal regions, most probably to the Oslofjord in the Southeast. There are only 27 pieces of rhyolite, but it does show contact with the west coast, most
probably with the rhyolite quarry at Siggjo at Bomlo in Hordaland. Rhyolite is also a possible chronological marker, as the type of rhyolite found in Hardangervidda is dated to ca. 6000 – 4000 BP on coastal sites (Indrelid 1994, 172, 275-82; Solheim 2010). The raw materials and technological traits supports the notion that both Eastern and Western groups have used the high mountains, making them an arena for social interaction between them (Solheim 2010).

Local raw material like rock crystal, quartz and quartzite are found as boulders and in outcrops in the mountains. One of the large quartzite outcrops in the Lærdal Mountains is also a very distinct feature in the landscape. Such outcrops with fine-grained quartzite or rock crystal may well have been among the focal points that structured movements in the landscape (Uleberg 2004). One of the oldest cairns in Hardangervidda may have been built to guide the way to a rock crystal outcrop. Similar cairns are used even today to guide the way through the landscape. Four peat-covered cairns have been found. C14-dating of the peat indicate that the cairn at the rock crystal site must be older than 2-300 BC, and the oldest cairn earlier than 7-800 BC (Sigmond 2008; Pedersen 2011).

Boulders and stones of quartz/quartzite can be found many places. Some sites are very small, less than 1m², but can contain several local materials treated in different ways, perhaps picked up close to the site (Uleberg 2007). At Gyrinosvatnet, local quartzite material from several sites has been refitted, reflecting movements to and from the sites around this lake (Schaller-Åhrberg 1990).

3.2 The Forest Limit

The forest-limit is a major environmental factor in these mountains. Today it is the result of a combination of tree felling for firewood and grazing by goats and sheep. Since the 18th century this activity has increased the effect of lower temperatures during the little Ice Age and kept the forest-limit down. Today, both the end to shieling and a temperature increase has resulted in a regrowth of trees at higher altitudes.

The Stone Age sites in Hardangervidda are today above the tree-limit. To understand the ecological and social space it is necessary to establish whether the sites were above or below the forest-limit, and whether this was a totally different environment or more or less a continuation of the forest at lower altitudes.

The Holocene climate history of Norway can be divided into three successive parts. The first phase is characterized by rather high temperatures, the melting of the last glaciers and plant immigrations. During the second phase, the Holocene thermal optimum, the pine forest reached its maximum as early as around 8700-8500 BP. The forest-limit varied along an East-West-gradient, and reached 1250 m asl. in the central area, 1200 m in the eastern parts and 1100 m asl. in the western parts of Hardangervidda. The third phase is characterized by a temperature decrease and consequently a forest-limit decline (Selsing 2010). Selsing advocates a beginning of the pine tree decrease from 6700 BP, then an expansion of birch but a further decline of the pine forest limit from 5500 BP and an increased decline in forest- and tree-limit from 4700 BP (Selsing 2010, 89, Table 11). Faarlund and Aas (1991), on the other hand, argue that the tree-line was rather constant from the early maximum until the climate deterioration at the beginning of the Iron Age (2500 cal. BP). There is both pollen and megafossil evidence of a birch forest belt at higher altitudes than the pine during the Holocene which may have reached a maximum of 1400 m asl. (Faarlund 1991; Aas 1999).

3.3 Big Game

Traditionally, the reindeer has been seen as very nearly the only game for the Stone Age hunters of the Norwegian high mountains. After
the Ice Age, the reindeer population in South Norway had a coastal adaptation. Gradually, the reindeer retracted from the coastal areas, and is today only found in the high mountains. The oldest occurrence of reindeer bones in Hardangervidda is dated to 8360 – 8100 BP (Indrelid 1994, 239). There are few finds of bone from the Stone Age in Hardangervidda, but reindeer occur at six sites, moose at one. Other osteological material is classified as reindeer/deer and unspecified mammals. A covariation between periods with worse climatic conditions for reindeer and fewer sites has been proposed, but is difficult to corroborate. The hunters in Hardangervidda have probably hunted all three types of big game; reindeer, moose and deer (Selsing 2010).

The Stone Age sites have generally been interpreted in light of reindeer hunting, since the Norwegian high mountains have been seen as a last refuge for the Palaeolithic reindeer hunters of Europe. The reindeer hunt in Northern Germany is very visible in the archaeological records. The osteological material found at Stellmoor show however that a wide range of species were represented (Krause 1943, 57-8), some of them living in forested areas. This indicates that the hunters had a wide subsistence base. It would be more in accordance with ethnographic parallels to think of Late Palaeolithic hunting groups not as highly specialized reindeer hunters but as hunters in exposed environments that have to use all available resources. These groups could be seen as marine hunters that hunted reindeer and moose as part of their annual cycle (Nordqvist 2009). This gives all the more reason to view also the first inhabitants of Norway as marine hunters with a wide resource base.

4. Social Space

Social space is created through interactions acted out against possibilities perceived in ecological space. The migration to Norway has been described as a gradual process where marine adaptation gradually became a more important element in the annual cycle (Welinder 1981). There seems, however, to be an initial hiatus just after the melting of the ice cap. When the sites first occur in the Early Preboreal, they are found all along the Norwegian coast. This indicates that there was a shift in technology and landscape perception that suddenly opened up new areas for habitation (Bjerck 1994; Naeroy 1999). This site distribution reflects the creation of a new ecological and social space.

4.1 High Mountain Sites and the Forest Limit

Today the Stone Age sites in the high mountains are in a treeless landscape, but the vegetation history indicates that most of the sites were not above the tree-line. Several of them were in the pine forest, and most of the others were among birch trees.

The importance of the tree line has been shown earlier in Lærdalsfjellene further south, were the prehistoric treeline coincides with the site distribution. All sites have been in either birch or pine forest (Uleberg 2004). The same pattern is visible at Hardangervidda.

Sites dated typologically or by radiocarbon were used in the analyses. Since the unit in the analyses is the event, both stray finds, single activity sites and sites with multiple visits were included. The material studied here shows that only 2 of 41 mesolithic sites are above 1240 m asl. and none above 1300 m asl. Out of 61 Neolithic sites 13 are above 1240 m and above 1300 m there are only 2 stray finds. As previously mentioned, the pine forest reached 1250 m asl. and the birch above may have reached 1400 m (Fig. 1).

On this background, the reindeer hunt does not seem to be the only reason for the distribution of sites in the high mountains. At Hardangervidda, there is nevertheless a positive correlation between sites and the present
reindeer spring, summer and autumn grounds, and hardly any sites in the Northeastern/Eastern part of the plain where the reindeer stay during winter (Pedersen 2011). This indicates that a seasonal reindeer hunt has been of importance, although it has been only one of several possibilities. The transition zone between two ecological zones generally gives more access to resources and the birch belt above the pine forest has been just that. The sites at this altitude have been a good choice for hunters of moose and deer in the forest and reindeer above the forest limit (Selsing 2010).

4.2 Landscape Attractors

The sites are mainly found around lakes. The material from Hardangervidda shows that the proximity to water is also valid for modern, artificially regulated lakes, which indicates that the co-variation is actually mainly dependent on the survey strategy. (Indrelid 1994, 219-22). The sites in the Lærdal mountains show affinity not only to lakes but also to large quartzite outcrops (Uleberg 2004). Both the landscape and survey strategy was different in the Nyset-Steggje-mountains. Larger lakes are not as dominant here as in the other two areas. Planned roads into the construction area were surveyed, leading to more surveying further away from the lakes. In addition, natural science results led to re-surveying of selected areas. As a result, the affinity between lakes and sites is not as strong at Nyset-Steggje as in other high mountain areas. An analysis of 76 dateable sites around natural lakes in Hardangervidda showed a close affinity between Neolithic sites and lakeshores whilst the Mesolithic sites do not show any particular pattern. Bronze Age sites however, seem to be further away from the lakes (Indrelid 1994, 221-2). However, when a major subsistence shift took place not at the transition to the Neolithic but at the beginning of the Late Neolithic, the Late Neolithic sites should be grouped together with the Bronze Age sites.

An analysis of 76 dateable sites around natural lakes in Hardangervidda showed a close affinity between Neolithic sites and lakeshores whilst the Mesolithic sites do not show any particular pattern. Bronze Age sites however, seem to be further away from the lakes (Indrelid 1994, 221-2). However, when a major subsistence shift took place not at the transition to the Neolithic but at the beginning of the Late Neolithic, the Late Neolithic sites should be grouped together with the Bronze Age sites.

The question is whether a shift in landscape perception at the onset of the Late Neolithic is detectable also in Hardangervidda. We assume there is equal chance for sites from all periods to be preserved and detected. An analysis based on georeferencing maps with the old shorelines would change the result, but sites that were submerged without survey and documentation would also bias the material. The present result shows the relative distribution of events from different periods documented through excavations, surveying and stray finds. Including the stray finds, less than one third of the Late Neolithic sites are within 25 m of a lake, as opposed to more than 50% of the older sites and over 60% of all Stone Age sites. This suggests there is a shift in the social and ecological space from the earlier parts of the Stone Age to the Late Neolithic.

4.3 Continuity or Change - a Shift in the Reading of the Landscape.

There are many indications of major changes at the transition to the Late Neolithic in Norway. In the coastal areas, there is stronger evidence of long distance contacts, and a pastoralist economy starts to dominate. In the Nyset Steggje high mountains, it is possible to see a shift in the site distribution. From then on, sites are located more in the same way as in the Iron Age and later periods. The sites are situated like the shieling sites, indicating that the shift from hunting and gathering to pastoralism is also visible in the high mountains. Also in the high mountains, the main transition occurred at the onset of the Late Neolithic (Prescott 1995b). This pattern can also be found in the site distribution in Lærdalsfjellene (Uleberg 2004) but not as clear as in Nyset-Steggje.
5. Single Site to Activity Area

In this study, the traditional archaeological site has been used as the main unit of analyses. When moving from single sites to activity areas, we approach a different understanding of the prehistoric landscape and how agents created and recreated environmental and social space through movement and activities. Many of the sites can only be broadly dated to the Long Stone Age but add to the understanding of movement. Lakes and rock outcrops have been focal points in the social space. Around the lakes, artefacts and debris can be found over longer stretches. The sites closer to lakes are generally larger and have been of more importance than the smaller sites at further distances. All sites are below the forest limit and varied osteological material indicates hunters with a wide resource base and not specialized reindeer hunters. Raw materials like rock crystal and quartzite indicate movements between the sites and local outcrops in the mountains, just as flint and rhyolite show contact with coastal areas. The high mountains may have been meeting places for groups from east and west.

This analysis has corroborated the hypothesis that environmental factors like the forest-limit and distance to water have been important factors for both hunters and pastoralists during the Long Stone Age. Events that can be placed in a narrower timespan suggest that a change to pastoralism influenced landscape perception also in this part of the high-mountains in South Norway.

References


Chalcolithic Territorial Patterns in Central Moldavia (Iaşi County, Romania)

Robin Brigand
UMR 8215 Trajectoires, France

Andrei Asândulesei
Alexandru Ioan Cuza University, Romania

Olivier Weller
CNRS UMR 8215 Trajectoires, France

Vasile Cotiugă
Alexandru Ioan Cuza University, Romania

Abstract:
This paper aims to compare spatial and temporal distributions of archaeological evidence in the western part of Iaşi County. Applying integrated approaches through GIS analysis, it aims to explore natural, economic, and social phenomena involved in territorial trajectory during Later Prehistory (4600-3500 BC). In the chronological framework of the Cucuteni culture, different kinds of spatial analysis are computed (viewshed analysis, kernel density estimation) in order to strengthen the control of the Bahluiet-Valea Oii basins, well-known for its fortified settlement density and its extremely suitable soils for agriculture.

Keywords:
Spatial Analysis, Cucuteni Culture, Viewshed, Kernel Density Estimation, Romania

Despite a long tradition of studies on Moldavian Neolithic and Chalcolithic cultures, the analysis of human communities’ territorial behaviour remains underexploited. This work combines concepts used in spatial archaeology with the potentiality of a Geographic Information System (GIS) in order to mobilise archaeological artefacts in a large-scale setting and multiple thematic scopes. The general goal is to evaluate how prehistoric territories are constituted and how natural resources were driving factors for these farming groups of eastern Romania. Visual analysis and spatial patterning allow us to describe territorial models which explain the original organisation of these territories.

1. Regional Setting

Located in the south-west area of the Moldavian Plain, the studied area covers the hydrographic basin of the Bahluiet, limited to the East by its confluence with the Bahlui. With the intention of defining a more restricted area according to the degree of advance of the archaeological map, the span of the study extends to the two basins of Bahluiet and Valea Oii as strictly defined by the outlet located in the downstream part of Sârca (Bâltați), at the merging of these watercourses (Fig. 1). Those watersheds, respectively of 300 and 95 km², differ by a vast interfluve (a landform composed of the relatively undissected upland between two adjacent valleys containing streams flowing in the same general direction). This is a cuesta landform, a ridge formed by gently tilted sedimentary rock in a homoclinal structure, slightly bulged with one long and gentle side (dip slope) conforming with the dip of the resistant beds that form it (towards the South and Bahluiet valley), and the other steep side (scarp slope) formed by the outcrop of resistant rocks (towards the North and Valea Oii valley).
Hydrography is the main factor of the current appearance of the Moldavian plain. Water has easily carved the geological sedimentary rock made of clay and sand. In the higher regions and on the Western and Southern limits, it has collided with sandstones and less crumbly sarmatian limestones of the Central Plateau or the Suceava Plateau. Above the marls and loess clay of the hydrographical basins of the Bahluieţ and Valea Oii, different types of soil have been observed. They belong to two main categories determined by the climatic zoning: the level of illuvial clay, made up of brown and grey steppic soils, that are occasionally found on the plateaus that limit the study area (West and South); the level of mollisols (chernozems) that occupy most of the Moldavian plain and its lower parts (Bacăuănu 1968). The former characterise surfaces that are currently covered by forest – or recently cleared – or by sylvosteppic forests. The latter include the back of cuestas and low interfluves, the terraces of the Bahluieţ and of the Valea Oii or the slightly steep sides, which are generally covered by meadows or fields.

2. Acharaeological Database

Diverse actors have, since the end of the 19th century, marked the prehistoric archaeology of the Eastern Carpathian Mountains, with important discoveries that are now part of the institutional historiography of the Cucuteni culture (e.g. Zaharia, Petrescu-Dîmboviţa and Zaharia, 1970; Monah and

Figure 1. Distribution map (A) and viewshed classification (B) in the Bahluieţ-Valea Oii watersheds during Cucuteni period (4600-3500 BC).
Cucoș 1985; Marinescu-Bîlcu 1993; Văleanu 2003; Boghian 2004; Bem 2007). The study area is particularly well-documented. Our study benefited from numerous geographic or thematic inventories that have been carried out at the scale of Moldavia, county or main geographical units. Given the complex history of research and an undeniable wealth of data, an exhaustive yet critical database has had to be achieved, gathering all available information: the context of the discovery, the chronological frame, the nature of the site, the cartographic and bibliographic observations, the accuracy of archaeological information, the field survey data, and the nature of georeferencing.

Excavation can provide a precise chronological framework from exhaustive samples. Dating a site from the single ceramic artefacts collected by field survey questions their very representativeness and reliability for periodic maps (Fig. 1A). The number of archaeological sites for Cucuteni A (4600-4100 BC) is the highest: 40 sites, among which 26 (34.5 %) were not occupied at a later date. As for the Cucuteni A-B, researchers have already noted the very low number of sites for this period. In our case, 8 sites are located in the Western area. This reflects the research difficulties (due to research conditions with the lack of abundant painted ceramic in field survey) rather than the retraction of the settlement. Since our study took a particular interest in long-term settlement patterns, it seemed appropriate to group together the Cucuteni A-B and Cucuteni B sites. Added to Cucuteni B, the number of sites considering a period stretching approximately from 4100 to 3500 BC reaches 33, among which 19 (25.5%) are not occupied during Cucuteni A. Fourteen sites (19%) are occupied from Cucuteni A to Cucuteni B. These are stable, generally significant economically and socially, and they attract settlement in the long term, over a thousand years.

The issue of archaeological classification has largely mobilised the scientific community. From the 1970s, the community started using topographical criteria in order to distinguish between different types of settlement. Looking at former inventories shows a more or less elaborate classification between higher, lower, or medium positions, but these are not always relevant given their variability depending on considered territories. Indeed, the appearance of territory plays an important part in the settlement patterns. For example, it has been observed that settlements sometimes selected steep slopes and cliffs formed by the outcrop of resistant rocks, such as the right side of the high and medium-high Valea Oii valley. Nevertheless, the topographical criterion on its own is insufficient to establish a valid hierarchy: it must necessarily be associated with other data, which facilitate the creation of a coherent hierarchical index.

The first criterion involves the presence or absence of man-made defensive structures. This points out social as well as spatial inequalities. In our area of study, seven sites are turned into defensive structures during the Cucuteni. This helps us to rank the data as follows. The lowest level is constituted of "occupations". These are small sites that only provided a limited number of ceramic remains and no obvious element of domestic architecture or materials of quality. They constitute an important category (28 sites, about 37.5%). This probably includes temporary sites, characterised by a strong mobility. Yet they are often little delineated and insufficiently surveyed. Contrary to these occupations, simple settlements display architectural structures and artefacts of quality (figurines, painted ceramics, bone and flint tools). Significant in terms of size and relatively numerous (21 sites, which constitutes 28%), they differ from hilltop settlements (19 sites, 25.5%), which are limited by steep slopes forming a headland open on one side. Low terrace settlements, closed on one side, are considered as simple and not hilltop settlements. Naturally, fortified hilltop settlements (7 sites, about 9%) are characterised by the existence of an man-made fortification. The generic category "settlement" includes sites with abundant artefacts and
house remains, generally stable, structuring the settlement pattern.

The georeferenced archaeological database depends on a precise protocol that need not be presented in detail here. Surveyed sites have been mapped by differential GPS (38 sites, more than 50% of the total) precisely where the strongest concentrations of ceramic remains have been found. When dealing with settlements well-defined by topography, the edges of the site have also been noted. The other sites were manually located using the descriptions contained in the inventories and the combined use of 1970 topographical maps at a 1:25,000 scale and maps dating from the first half of the 20th century at 1:20,000, as well as orthophotography. Indeed, among all the sites, the position of 21 of them (28%) can be placed with a margin of error of approximately 50 metres. Fifteen other sites (20%) are so-called imprecise locations, as the margin of error varies between 50 to 200 metres. Only one location remains inaccurate and located in the centre of the village.

Starting from this pattern of dots, a series of spatial analyses have been undertaken, benefiting from the wealth of specialised literature, a French project developing a model for spatial processes (Gandini, Favory and Nuninger 2012), and several experiments carried out in Neamţ County (Weller et al. 2007; Weller et al. 2011).

3. **Viewshed Analysis**

In this section, the study of settlement patterns uses several parameters offered by GIS: viewshed analysis (Fig. 2), density estimation (Fig. 3) and more broadly, the anisotropic travel-times (Fig. 5). A major limit must be stated: since these studies are mostly based on indexes defined by field surveys, it is impossible to study their contemporaneity more precisely than in broad archaeological phases corresponding to 500-600 year periods (respectively Cucuteni A and Cucuteni A - B and B).
Viewshed analysis is one of the classic tools offered by GIS and has thus been largely used to resolve issues of territoriality peculiar to human societies (Wheathley and Gillings 2002, 202-216; Conolly and Lake 2006, 225-233). They allow the rephrasing of some crucial notions concerning the study of the forms of settlement: territory, considered as a transformed, occupied and appropriated space socially controlled by a group; status or rank, which corresponds to the different levels of hierarchy of the settlement; relations between the different parts of a spatial system, that is, the issues of specialisation and synergy of archaeological entities. Visibility calculation determines areas that can theoretically be seen from different observation points. Three essential shortcomings must be raised. The first one deals with the DEM resolution, since the calculated viewshed results depend on its accuracy. In our case, the small pixel value (25 m) elaborated by K. Ostir of Ljubljana University (ZRC Sazu) from ERS radar images allows for accurate and precise results. The second limit depends on how visibility analysis is programmed in each particular GIS software package. The software we used – ArcMap and Erdas Imagine – does not allow for any choice in the way in which visibility analysis is computed. With one single set of data, different algorithms produced different results. Several tests dealing with field observations resulted in our confidence in one of them: the Leica software package (Erdas Imagine). The third limit is fundamental and arises from the weakness of paleoecological data, such as vegetation and tree cover. Though forests can have a decisive impact on the field of vision, this study ignored this parameter because reconstructing vegetation history presented so many difficulties.

Several analyses have been carried out, from the theoretical assumption of an observer whose height is 1.7 m above ground, according to a standard offset. The field of vision is limited to 12 km, according to field observation (in different weather condition) and ethnographic information. It also corresponds to one day walking round trip. Besides, this value approaches general settings used in archaeological analysis. So, this paper assumes that a village, a small group of domestic units, or cattle located in an open landscape, are visible at 12 km in favourable weather conditions. The simplest way of visibility calculation is a binary map distinguishing between visible or invisible target cells from a specified viewpoint. The visible spectrum can then be quantified in square kilometres. Its classification using standard deviation offers a first level of hierarchisation according to the importance of theoretical visibilities for each site considered (Fig. 1B). The visibility map might be associated to one or more viewshed maps. The result is a multiple viewshed map in which the values are either 1 (visible) or 0 (not visible). Each map cells is noted 1 if it is visible from at least one viewpoint. On the other hand, the map algebraic sum of two or more binary single viewshed maps creates a cumulative viewshed (Fig. 2). Then, the cell values are integrated ranging from zero to theoretical maximum of the number of viewpoints, although this will only occur if at least one cell is visible from all viewpoints. The field of view being defined, the maximum value generally cannot be equal to the number of archaeological sites. This method is also used for defining a qualitative index of visibility taking into account not several archaeological sites, but the whole set of points lining a given settlement. It allows to qualify the visible spectrum that gives an account of the different viewpoints whether the observer stands at the centre or on the side of the settlement (Fig. 5). Contrary to cumulative viewsheds, the map of multiple viewsheds results from the association of several simple or binary visibilities. The outcome is thus a map of visibilities in which the values equal 0 or 1 (meaning the pixel is visible to 1 observer at least). Subtracting multiple visibilities for Cucuteni A-B and B from Cucuteni A allows us to identify the dynamics of seen and unseen areas (Fig. 4b).
4. Density Estimation

The kernel density estimation (KDE) provides an estimation of the density of a point pattern. For a circular kernel estimation, the density value obtained takes into account the size of the neighbourhood: thus, an area surrounded by other high-density spaces will in turn become more dense. The assigned weight decreases proportionally to the distance from the centre of the window. This method is well known since the 80's (Silverman 1986) and has largely been used for archaeological applications for intra-site or inter-sites analysis (Baxter, Beardah and Wright, 1997; Nuninger et al. 2012). The density estimations depend on two parameters: \( k \), the kernel function chosen; \( h \), the radius chosen. ArcGIS uses a quadratic kernel function, with no alternative choice. In archaeological analyses, the choice of the radius (\( h \)) is the main parameter, for it determines the smoothing of the data. Generally, using too small radius will produce irregular surfaces, similar to a pattern of dots. On the contrary, too large radius will result in a loss of accuracy, favouring general trends and preventing the observation of settlement patterns.

For determining the best radius, this study uses a graphic approach inspired by the ArchaeDyn programme (Nuninger et al. 2012, 32) and applied in the Neamţ County (Weller et al. 2011). It sets a curve of the maximum values obtained according to a series of calculations linked to a given interval (200 m). The inflection point of the curve corresponds to our situation, estimated at 900 m (in reality between 800 m and 1000 m). A major limitation is raised when the KDE method must be applied to archaeological data. In order to calculate the settlement densities by period (Cucuteni A, Cucuteni A-B and B), the sites with chronological dating that has been imprecisely attributed to the Cucuteni culture are not taken into account. In order to overcome this bias, a weighting according to the length of each period has been made. In the database, a site that can undisputedly be attributed to a period has a value of 1, and a site that undisputedly does not belong to a period has a value of 0. Each imprecisely dated site is given a value of 0.45 for Cucuteni A and of 0.55 for Cucuteni A-B and B. The main advantage of this method is that it takes into account sites that were previously excluded from the analyses by focusing on precise chronological periods. The density

Figure 3. Cucuteni sites density with chronological and quality weighting factors.
varies with a lesser weight for these sites in order to give a more realistic image (Fig. 3).

Concerning the nature of the archaeological site, an arbitrary weighting is introduced on a scale of 1 to 4, which allows to discriminate quality settlements. Since this study tries to establish the importance of hilltop and fortified settlements in the organisation of Chalcolithic territories, those reaching a 4 should be given priority. They were most probably richer in terms of population and power, while the weight of short-lived or undefined occupations should be attenuated. Thus, small settlements, usually identified by fieldwalking surveys, are attributed a value of 1, whereas simple, hilltop and fortified settlements range from 2 to 4. The product of the two weighting factors, i.e. the nature of the site and its chronological framework, allows us to define a value which will be used for these KED analyses. Thus, simple settlements of which dating is uncertain are attributed to the Cucuteni but without specifications, will be given a value of 2x0.45 for Cucuteni A and 2x0.55 for Cucuteni A-B and B (Fig. 3).

With those density maps that provide a broad view of settlement processes by chronocultural period for the Bahluiţ-Valea Oii hydrographic basin, differential density maps have been associated in order to visualise positive and negative evolutions between the two chronological sequences. The instability index, whether it be negative (abandonment) or positive (creation or development) is obtained by subtraction of the site density, weighted by its nature and chronology (Fig. 4). Negative values correspond to deserted sites; conversely, positive values correspond to created sites (i.e. new site or rise of the site's status). Finally, while a value 0 indicates the absence of occupancy between the two periods, it also indicates the stability of the settlement throughout the two periods. Hence, in order to differentiate between the two parameters, the location of stable sites between Cucuteni A, A-B and B has been specified (Fig. 4a).
5. Results and Discussion

Spatial analyses are mainly based on distribution maps from which graphic models are elaborated. In order to fully understand the method used in this study, two essential biases must be stated. On the one hand, archaeological information cannot be exhaustive and is bound to be partial. On the other hand, it is impossible to prove the contemporaneity of several sites placed in a centuries-old chronocultural phase (except in specific cases). The issue, while hard to solve in absence of radiocarbon dating and detailed excavations, can nevertheless be studied in terms of settlement patterns.

5.1. Regional Distribution

A glance at the maps showing the archaeological spatial organisation is sufficient to define the general characteristics of their geographical distribution: settlements are tightly linked to the stream channels since they systematically stand on the edges of alluvial or erosive terraces, as well as on the ridges of cuesta landforms lining the watercourses. A few exceptions can be singled out. First of all, the Southern part (Fig. 1) and more precisely the area where the tributaries of the right bank are gathered, a looser settlement pattern has been observed. It does not reach the concentration numbers of other settlements in the Valea Oii and Bahluieţ valleys. The settlements are distributed along the Ciunca and the Albeşti and, to a lesser extent, along their respective tributaries. Most Southern sites seem relatively isolated. They can be found in higher parts of rather minor and probably seasonal watercourses.

How can the unusual occupation in this area be explained? The topographical variable and the socio-economic environment suggest a few hypotheses. The morphology of the territory, with numerous narrow and symmetrical valleys with relatively steep slopes whose summits reach 150 to 200 m, does not provide a favourable place to settle. Furthermore, forest coverage may have been more than the current situation. In this context, the settlement pattern might be derived from pioneer settlers. It is suggested by isolated settlements, dissociated from the settlement patterns of the Bahluieţ valley. Two arguments suport this idea. First, the original topography offers many limited viewsheds, well-established by the map of hierarchies according to potential visibilities (Fig. 1B), as well as low viewshed competition emphasized by the cumulative viewshed maps (Fig. 2). The second argument points out the temporary nature of this occupation, without stable or fortified settlements (Fig. 1a).

On the contrary, in the Northern part (Valea Oii and Bahluieţ valleys), settlement patterns change radically. The archaeological distribution is very dense and organised according to several centres of population, usually indicated by the presence of stable settlements established on cuesta ridges or low alluvial terraces allowing them to control the fluvial landscape. We should keep in mind that the landscape is more open due to the rather flat topography, probably stimulating new occupations with close intervisibility.

5.2. Settlement Pattern

The change between Cucuteni A and Cucuteni A-B and B is well-established thanks to the combined use of distribution maps and density analyses (Fig. 1a, Fig. 3). A regression can be noted during Cucuteni A-B and B, characterised by a noticeable decrease in the number of settlements. This suggests a phenomenon already observed in the foothills of Neamţ County: a retraction and concentration of settlement according to specific choices about territorial control and land resources. Between Cucuteni A and Cucuteni A-B and B (Fig. 4a), the number of abandoned sites is high, notably in the middle-low Valea Oii valley. None of the eight sites seems to continue during the next period. This settlement duration probably reflects an agrarian colonization, as confirmed by the dating of at least three of them to
Cucuteni A3 (c. 4350/4300-4150 BC) (Fig. 5b). The creation of sites during Cucuteni A-B and B is also important, in most cases reaching new adjacent areas. It implies, after the extension of Cucuteni A and the massive diffusion of its settlements, that a displacement and resettlement of the same populations occurred. New creations of sites are often found in direct proximity to deserted sites. Thus, they indicate a displacement and rarely, the occupation of new land.

Next to these displacements of population between Cucuteni A and Cucuteni A-B and B, another situation should be documented, which is less a displacement than a genuine strengthening of the settlement pattern. The network is not radically changed, for it relies on a former pattern unchanged by new occupations. Except the desertions in the Valea Oii valley and in the Southern part, density maps (Fig. 3) show a strengthening of previous polarities, whether in the upper and middle Bahluiet valley, more densely occupied, or in the upper Valea Oii valley. In this area, it has been clearly observed that following a period of extension and diffusion, a resettlement occurred around the main fortified sites (Cucuteni-Cetăţuie, Cucuteni-Dâmblu Mort, Stroeşti-Pietrârie). This new territorial organisation seems to characterise the final phase of the Chalcolithic, as the dynamic viewshed map also records (Fig. 4B). Indeed, apart from the drastic reduction of visible spectrums in the middle and lower Valea Oii valley, this map highlights stability of visible areas.

5.3. Territory, Mobility and Specialisation

The territory is defined as an area socially appropriated and domesticated by one or several communities, in which a population
exercises immediate leverage for its activities of production or hunting in relation to the ecological context (field, pasture, forest, etc.). For instance, some fortified settlements, stable over a few generations and evenly distributed along the main fluvial corridors, were central places tied to its economic activities but also to probable affective factors. Next to stable and federating entities, temporary and mobile settlements (small farmhouse, agrarian annex, sheepfold, etc.) belonging to a specific socioeconomic process.

Settlement patterns consist of a dense network, with a strong hierarchy conveyed by a variety of archaeological sites, whether fortified settlements, hilltop settlements without fortifications, open settlements or temporary occupations. Hilltop settlements, sometimes fortified, can generally be found on a side of cuesta or a high terrace of ridges. It is limited by cliffs and, in some cases, by an open defensive ditch. Located high up and in an overhanging position, it visually controls a territory of dozens of square kilometres, and thus shares intervisibility with other settlements, but also with more minor sites established in variable topographical contexts. Different hierarchical levels have been observed in such cases. Fortified settlements, appearing as federating centres, have a very different access to resources than settlements located in alluvial plain or on a low terrace next to a watercourse. Taking soil information into account, one may suppose that hilltops are made of areas exploited for wood and breeding purposes. In the valleys, these settlements are located near pastures and water resources, thus probably intended for agricultural and pastoral purposes. Many small farms, probably seasonal and mobile according to economic production necessities, surround these federating centres.

Viewshed analysis provides more information. The example of the Valea Oii valley can be studied here for short periods (Cucuteni A2: c. 4500-4350/4300 BC; Cucuteni A3: c. 4350/4300-4150 BC) (Fig. 5). For Cucuteni A2 (Fig. 5A) 8 sites are within the viewshed of Cucuteni-Cetățuie fortified settlement, whether they are important settlements, fortified or without fortifications, or small agricultural units (generally imprecisely dated within Cucuteni A). Compared to Cucuteni A3 (Fig. 5B), with the creation of two hilltop settlements (Filiași-Dealul Mare and Podișu-Dealul Boghiu-Crescătorie), the analysis reveals a colonisation of the low valley. The remarkable concentration of sites under Cetățuie throughout Cucuteni A2 can be viewed as the expression of the occupation of a territorial unit, less than one hour away from Cucuteni-Cetățuie, where several settlements would have held different functions and coordinated the diffusion of satellite sites. A similar settlement association has been found in the low Valea Oii valley during the Cucuteni A3. One fortified settlement (Filiași-Dealul Mare) is directly connected to other sites:

- the nearby site of Filiași-South West Dealul Mare that, even though it is fortified, is still close in the alluvial valley and thus benefits from a specific access to resources;
- the more distant site of Bâlțați-Dealul Mândra, located on a watercourse and in an open landscape, was probably based on agricultural production and/or hunting.
This synergy between sites with access to specific resources (water, soil and wood resources) defines a geographical and economic territory, in other words an area of land appropriated and exploited by a human community.

In the Valea Oii valley, it is particularly interesting to note that the two territories – the first one is structured on Cucuteni-Cetățuie, the second one on Filiași-Dealul Mare – are over 10 km apart, which is more than a two-hour walk (Fig. 5). The site of Balș-Bejeneasa, although imprecisely dated in the Cucuteni A, is set in an interesting location since it is halfway between the two settlements.

This settlement pattern can also be found in other zones of our study area, such as upstream of the confluence of the Bahluiet and Ciunca rivers. In this case, the strong visual competition between the settlements illustrated on the panoramic photograph (Fig. 6) could reflect their mobility on both sides of the fertile Bahlueț valley without identifying any federating settlement.

6. Conclusions

This study has highlighted methods and problems in the study of settlement patterns in the 5th and 4th millennia BC. In Moldavia, similar settlement patterns have been observed in an earlier period, during the Precucuteni culture. They largely spread during Cucuteni A. In fact, Precucuteni settlement is distributed in the middle Bahluiet valley and to a lesser extent in the upper Valea Oii valley, where settlement gradually becomes denser during Cucuteni. In Cucuteni A, the number of settlements increases. Their wide distribution involves three simultaneous factors: first, the demographic increase provides new agents of settlement; second, the development of agricultural practices leads to a greater mobility; last, the intensification of territorial hierarchy leads to the emergence of federating centres of settlement. GIS analyses, whether dealing with visibility or density, reinforce these hypotheses while underlining the short-lived nature of these settlement patterns.

Acknowledgements

This work was made possible with the financial support of the Romanian Ministry of Education, Research, Youth and Sport: Sectoral Operational Programme for Human Resources Development 2007-2013, co-financed by the European Social Fund, under the project number POSDRU/89/1.5/S/63663 and POSDRU/89/1.5/S/47646 and the Exploratory research project PN-II-ID-PCE-2011-3-0825, The Ethnoarchaeology of Salt Springs and Salt Mountains of the extra-Carpathian zone of Romania, no. 219/05.10.2011 (Project Manager: Dr. Marius Alexianu). This study has also been funded by the French Foreign Office (Commission for Archaeological Research) since 2004.

References


Settlement Patterns in Drahany Uplands (Czech Republic): GIS and Quantitative Methods Based Approach

Lukáš Holata
University of West Bohemia in Pilsen, Czech Republic

Abstract:
Previous extensive field survey in the Drahany Uplands documented over 60 deserted medieval villages together with their fields. The (almost) complete medieval settlement network is used for further analysis based on GIS and quantitative methods. This research primarily aims to understand the reasons behind the abandonment of villages. Firstly, the regularities in the lay-out of villages and their fields are obtained. The probable motivations behind the choice of space for settlement areas are then briefly discussed. Some of spatial characteristics of the deserted fields are used in the reconstruction of the most intensely settled landscape at the beginning of the 14th century. Finally, some factors that contributed to village abandonment are mentioned.

Keywords:
Deserted villages, Field systems, GIS, Quantitative methods, Middle Ages

1. Introduction

The Drahany Uplands is considered the largest region with the most complete medieval rural settlements in the Czech Republic. A long-term (from 1953 to 1992) research project was undertaken here by Ervín Černý (especially 1979; 1992), a physician and also an amateur archaeologist, who located and documented more than sixty deserted villages preserved mostly in the woodlands of the region. Some of the villages are not mentioned in the written sources. Most of them were also documented together with the remains of fields which were rarely detected in other surveyed regions (cf. Klášť 1978; Smetánka and Klášť 1981; Krajíč 1980; 1983; 1987; Vařeka and Rožmberský et al. 2009; Vařeka et al. 2011). Therefore Černý’s outcomes represents a unique source of information, and it allows archaeologists to address a number of questions connected with medieval rural settlement which are outside the scope of more traditional methods of research.

The current research primarily aims to understand the motivations behind the abandonment of villages. The systematic study of the late medieval settlement changes is currently not a popular topic in the Czech Republic, and in discussions has largely predominated the conventional interpretation of these processes (cf. Nekuda 1961; Měřinský 2008). In this paper I introduce the preliminary results of this research. Furthermore, I aim to understand the past human motivations behind the choice of space for settlement areas and also to reveal what the densely populated landscape looked like at the beginning of the 14th century (e.g. Černý 1992, 116).

2. The Region – Natural Environment and Settlement Development

The Drahany Uplands are located north of Brno in Moravia (Fig. 1). The extent of my research region covers the area of 787 km² and more than half of it is represented by woodlands. Fertile lowlands can be found in the northwest, west and south parts of the region (minimum altitude – 238 m); the central, north and east parts are more hilly (maximum altitude – 735 m) with a higher proportion of barren soils (cf. Černý 1992, 9; Tomášek 2003). A major
watercourse is found along the western border of the region; smaller streams drain the rest of the area.

The whole region was colonized by Germanic inhabitants from Lower Austria from the mid-13th century onwards (e.g. Doležel 2003, 129). However, there is evidence for older settlement dated to the early Middle Ages and several insulated enclaves of settlement from the 9th century (Belcredi 1983, 40; Černý 1992, 12, 116; Doležel 2003, 135). The highest settlement density probably represents the beginning of the 14th century. Approximately at the same time the first abandonment of villages began not long after their establishment. However, the main peak of abandonment came at the first half of the 15th century (almost half of the deserted villages) and continued in a smaller extent until the beginning of the 16th century (Černý 1992, 127–128). During the 18th century eight villages were repopulated and new villages were founded (cf. Hosák and Šrámek 1970; 1980; Fig. 2).

3. Evaluation of Previous Research and Methodology

Unfortunately, Černý did not record the coordinates of the deserted medieval villages and their field systems. The only solution was to scan the plans from his concluding publication (Černý 1992). These plans were subsequently imported to ArcGIS 10 where they were georeferenced using constant elements in landscape (e.g. rivers confluence and meanders, peaks of hills, also crossroads etc.).

Afterwards, the plans were vectorised and the position data of villages and fields were obtained (Fig. 3). Possibilities of field confirmation are markedly limited by recent activities in the woodlands and bad

3 Evidences about six deserted villages in this period exist (mainly on the basis of the concentration of ceramics dated to first half of 14th century). These are not necessarily mentioned in the written sources, their features are difficult to distinguish and their total number might be higher.

4 The accurate duration is not known for either deserted village. In most of cases, they are dated by archaeological material. Written sources were also used, but these are not frequent. Therefore, dating is vague and except the first and the main phase it is not possible to distinguish other significant periods of abandonment.

5 Although he did not mention the coordinates, his field measurings were very accurate.

6 I have chosen only these passpoints which correspond to each other in the current and Černý’s maps. Nevertheless, small divergences could be included (max. several metres). The accuracy of particular fields is mutually validated because close situated fields occur in a number of Černý’s plans. Slight mistakes cannot expressively influence the general patterns.
preservation of the majority of field systems. Data about newly-documented villages and other field surveys which specified the dating and location of deserted villages were extracted from other sources (especially Doležel 1993; 1999a; 1999b; 2000; Doležel and Černý 1999; Doležel and Plch 1999; 2000; Režný 2001). The position data of living villages were gained from II. Military Mapping (1836 – 1852) in Web Map Service (WMS – Geoportal 2012), which recorded the original location of settlements and their extent; basic characteristics were gained from other sources (especially Hosák and Šrámek 1970; 1980; Doležel 1996).

The digitized contour map at a scale of 1:10,000 (ZABAGED) was interpolated to DEM and other specific rasters (especially slope, aspect, cost distance⁷, stream order and network, fuzzy membership and overlay⁸) were derived. Soil types were determined according to WMS (Geoportal 2012). Subsequently, the spatial characteristics of deserted medieval villages and their fields were analysed in ArcGIS 10 and they were statistically tested in STATISTICA 10. It enabled us to highlight regularities in the lay-out of villages and their field systems and settlement patterns.

### 4. Choice of Space for Settlement Areas

I consider the understanding of past human thinking and behaviour as one of the crucial objectives in research on the abandonment process. It is evident that this topic is very extensive. In this paper, I try to detect some basic environmental characteristics of settlement areas, settlement distribution and the reconstruction of the past landscape including factors that influenced the choice of space for settlement areas and participated in desertion of some villages as well.

I suppose that the deserted fields alone may not reflect the whole picture because of their specific characteristics, hence a test set of fields was also created. These include the areas where fields are hypothetically expected⁹. Environmental qualities of both categories were compared with the whole region.

Rural areas were found in lowlands as well as highlands (262 – 683 m); inner 50% of values are between 403 – 561 m. Villages were generally located in lower positions in the centre of their hinterlands (fields, mining areas etc.). Although fields were positioned near the villages, some parts did not have to be easily accessible because of their longer distance from village or their placing on steeper slopes.

Every village was situated near a stream even if not all villages are directly adjacent to them (stream confluence – 25% of villages, along stream – 37%, spring area – 38%, not adjacent – 6%). Distance of villages from each

---

⁷ Although I am familiar with the frequently used Tobler’s formula (Tobler 1993), I have decided on a different solution algorithm for cost distance of village’s hinterlands, especially fields. It is clear that the speed of motion is influenced by a number of factors (slope, land-use, vegetation, season and weather conditions, etc.) and varies for each individual. Moreover, the medieval relief was probably different because of erosion. Therefore, a more benevolent algorithm was chosen, which incorporates the classes of slopes in the terrain.

⁸ Fuzzy logic was used for prediction of forested and deforested areas. Specific digital models were reclassified according to Membership types (functions) and threshold values to new models. Subsequently, they were combined together (including binary maps) based on overlay type Fuzzy And.

⁹ Around living villages were created 1 km buffer zones and current woodlands were clipped. The boundary between villages located in close proximity was measured in half distance as with Thiessen polygons.
other amounts 497 – 2544m; the median is 1470m. It seems logical that quality soils were preferred but only the northwest, west and south parts of the region are really fertile. In this way villagers had to use also less fertile soils and most of fields were situated on poor soils. However, settlements tend to be place close to soil type boundaries (64% of all hinterlands are directly intersected by this boundary).

There is a manifested tendency to choose lower slopes for villages (62% villages to 5° but their some parts intervened over 20°) and for their fields, too (preference for 0° – 8° slopes, some parts though reach a higher slope, even around 30°). The eastern slope orientation with variations to northeast or southeast significantly dominates in villages’ positions (47% of villages), whereas south slope orientations prevail in the region. I cannot explain this pattern as yet. There has been find a slight trend for the use of the south and southwest (as these get most light) slopes for fields. An avoidance of north slopes is evident for villages and less so for fields (with expansion to northeast and northwest).

The significance of settlement areas distribution in some aspects (altitude, slope and exposition of terrain) was tested by the chi-square test. It has shown that all null hypotheses (exploitation of expected conditions in region) cannot be rejected. It is possible to say that the region offered some more favourable areas for settlement, which however were not exploited. Thus, the choice of settlement areas could be also influenced by other factors (e.g. political, and social-economical; cf. Neustupný 1994).

Nevertheless, I attempted to detect the significant mutual combinations of certain chosen environmental characteristics10 which influenced the choice of space for settlement areas. I applied the Principal Components Analysis (PCA; after Neustupný 1997; 2005) for deserted ploughlands and the test set of fields. Two significant factors were extracted from the correlation matrix, they were rotated by Varimax method and the major regularities were gained. The first include altitude, fertile soils and border of soil types. The second include slope of terrain, exposure to the south and good accessibility of hinterlands. These results offer a hypothesis that the most desired areas were in lowlands, on fertile soils, with various soil types that guaranteed the best agricultural conditions. However, these possibilities are spatially restricted in the Drahany Uplands, and they were compensated in other settlement areas by mild slope areas which were easily accessible and also with dominant south exposures. However, many settlement areas still do not fall within these structures; therefore it is possible to suppose different motivation in choice of space for these settlement areas.

5. Reconstruction of Past Landscapes

Previously detected patterns can be used in the reconstruction of landscape with the highest density of settlement at the beginning of the 14th century (Černý 1992, 116). Existence of 143 settlement sites is assumed for this period whereas 101 settlement units are in the region at present. The results of the prediction can be used in future research of Drahany Uplands. For a clearer idea about the medieval situation, more specific spatial analyses can be undertaken and other factors that possibly influenced the abandonment process can be examined. Especially, the impact of human activities on natural environment would be included.

The Fuzzy and Boolean logic together with map algebra were used for the prediction. The first step determined the extent and ratio of forested and deforested areas. It resulted from the patterns of deserted fields; the slope and exposure of terrain and cost distance

10 The descriptors chosen were: portion of fertile soil in hinterlands (%), distance from soil types border (m), middle altitude of hinterlands (m), portion of territory with 3°slope in hinterlands (%), portion of south exposition (112,5° – 227,5°) in hinterlands (%), portion of optimally accessible terrain from villages (%).
from villages were involved\(^1\). However, some deserted fields were also documented by the field survey to be in unsuitable positions (e.g. steep slope). Therefore, the value of 1 was ascribed to deserted field polygons as well as to villages because these areas were certainly deforested. Consequent fuzzy overlay reflected the areas where the forested (0 – 0.6/0.7 values) and deforested areas (0.6/0.7 – 1 values) can be expected.

Finally, the areas with probable occurrence of meadows were extracted from the deforested areas. They could have been positioned along larger streams in alluviums and especially on fluvial soils (Fig. 5).

6. Abandonment of Villages

It is beyond doubt that war events together with general economic and demographic crises were significant factors in the abandonment process. However, no direct evidences testify for these factors in the Drahany Uplands\(^2\). Therefore, my approach tries to observe the factors in background which could contribute this process and considers the abandonment of settlement as a selective process. Villages with certain attributes or with their combination

---

\(^1\) Slope – mild slopes were preferred (Fuzzy Small), Aspect – south exposition was preferred (Fuzzy Gaussian), cost distance – profitable availability of hinterlands was preferred (Fuzzy Small)

\(^2\) This situation is especially caused by less numerous written sources. Although, there were archaeologically documented abandonment layers with charcoals and daubs, it is not possible to link them with particular units.
were the weakest elements in the settlement network and, therefore, they were more predisposed to abandonment (cf. Zölitz 1984, 36; Lewis, Mitchel-Fox and Dyer 1997, 152). I attempt to specify the role of the above mentioned characteristics in this process during the 15th and the beginning of the 16th century. I also try to find out whether deserted fields were different from the test set of fields in placement and utilization of the landscape. The Kolmogorov – Smirnov test (K – S test) was applied which can be used for all selected descriptors.

Generally, villages were abandoned more often in higher altitudes. Only few deserted villages have been found in lower altitudes (up to 400 m; the rate with living villages is 12.5% vs. 87.5%), conversely they predominated in a higher altitude (over 600 m; the rate – 61% vs. 39%) with least favourable conditions for settlement (cf. Černý 1992, 128).

Villages were also abandoned in areas with fewer major water sources. Deserted villages are slightly more often situated along smaller streams (71% of them vs. 47% of living villages) and in the spring areas (45% of deserted villages, 38% along the streams, 16% at the confluence) while living villages have more numerous occurrences along larger rivers (53% of them vs. 29% of deserted villages) and by the confluences (32% of living villages, 35% along streams, 33% in the spring areas).

Only 4 cases of deserted villages were documented in the most fertile soils. However, each of them was situated in the immediate vicinity of a town and in addition, their fields continued to be used up to the present. There is no case of a completely deserted settlement area on high-quality soils. On the contrary, there is evident tendency in abandonment of villages in one less fertile soil type (Fig. 6). Agricultural possibilities of these villages might have been very limited and that is why they were not very profitable for their landlords.

In comparison with deserted fields there appears to be another trend. Deserted settlement areas with smaller portion of mild slopes of terrain (up to 3°) include slightly higher portion of the south orientation of slopes. Test set of fields differs from deserted fields in all examined characteristics (as K – S test showed) and it has better conditions for agricultural exploitation.

In addition, deserted villages are located in areas with overall higher density of settlement (average distance for deserted villages = 1254 m vs. for living villages = 1523 m). However, there are cases that some of the deserted villages were located in an isolated position from other villages because of their surrounding relief (Fig. 7). This situation could have significantly limited or made more difficult the social contacts of the villagers. Similarly, some villages had difficult access to their hinterlands. Some parts of their fields could even be better accessible from neighbouring villages. Both factors, however, were not regarded as significant in the abandonment process in Czech archaeology.

Results of the spatial analyses demonstrate that deserted villages with their hinterlands were situated in worse conditions for settlement and they could have had a minor economic potential. Therefore, one enunciation suggests that at least some villages were established in bad positions for settlement. It is impossible to disprove this statement. Nevertheless, I am inclined to support the idea that the choice of space for settlement areas did not represent a poor choice and the inadequacies of these places came to fruition later during their exploitation. They could be dramatically accented during crisis events (e.g. wars) and marginal character of these settlement areas, caused by one or combinations of their characteristics could stand out. During the period with lower number of inhabitants (after war events), villages tend to founded in more advantageous conditions, 14 This refers to villages deserted from 15th century. The earlier desoration could be influenced by other processes connected with medieval transformation of settlement. It came out by the transformation of settlement structure which caused the concentration of villages.
which were also more profitable for landlords. These statements and this model still have to be tested by future researches. The archaeological sources will have to be combined with historical sources. Especially social-economic variables need to be incorporated to this research.

7. Conclusions

The research presented here draws upon previous extensive field survey of Drahany Uplands. Over 60 deserted villages were documented together with the remains of their fields. This almost complete medieval settlement network was analysed and evaluated in GIS and with quantitative methods. The basic settlement patterns were then revealed. The choice of space for settlement areas and motivations behind it were discussed: There was a tendency to settle fertile areas in lowlands and where it was not possible, settlement shifted to a flatter terrain with south exposure. For other positions different motivations are assumed.

Subsequently, the medieval landscape was reconstructed and the proportion of deforested areas was determined. This was about 8 – 16% higher than today. Finally, some factors which could participate in the abandonment process of these settlements were mentioned. It appears that the deserted villages were in less advantageous locations than the villages that survived. Disadvantages could have increased especially after war events.

These preliminary results have to be tested by future research. The main attention should be paid to a complex study of the abandonment process. The results of medieval landscape prediction can be utilized and environmental factors have to be combined with social-economic factors. This could represent the solution to obtaining the summary of attributes and their specific combinations that influenced the abandonment process and in this way the motivations behind it can be revealed.

Acknowledgements

This was supported by the project “The strategy of archaeological research in Europe”
I would particularly like to thank Lucie Čulíková and Monika Baumanová for their help during the writing of this paper.

References


Vlastivědný věstník moravský 53: 391–392.


Rural Life in Protohistoric Italy: Using Integrated Spatial Data to Explore Protohistoric Settlement in the Sibaritide

Kayt Armstrong and Martijn van Leusen
Groningen Institute of Archaeology, the Netherlands

Abstract:
This paper introduces a research project that employs a complex set of spatial data to examine sites primarily identified from field walking over a ten year period in northern Calabria (Italy). Our core aims are to understand both what these protohistoric pottery scatters represent in archaeological terms, and the processes by which they form and are discovered in the present. We describe the problems we have encountered with integrating very diverse datasets, across differing scales, and disciplinary cultures and languages. We give examples to illustrate the types of solution we have arrived at, and the impact this has on our understanding of sites known previously only as surface ceramic scatters.

Keywords:
GIS, Field Walking, Geophysics, Geomorphology, Integration

1. Introduction

Our project combines new information from different types of geophysical survey, excavation, intensive field walking and geoarchaeological and geomorphological studies with an existing body of knowledge containing LiDAR data, current and historical aerial photography, current and historical maps, landscape classifications, geological and cadastral data, as well as historical archaeological site inventories. We briefly present technical problems and solutions relating to the spatial integration of such data sets, particularly the steps necessary beyond simply displaying them within the same projection and coordinate system. We then focus on problems in the integrated analysis of archaeological heritage, remote sensed and geophysical data, which require a specialist understanding of the way the data has been obtained and processed. The problems of utilising historical data such as aerial photographs from the 1940’s and 1950’s and topographical survey information from the 1960’s will be considered.

One of the benefits of this integration process are enriched site descriptions that place ceramic sites within their landscape and historical context for heritage management purposes, since the addition of geophysical and geoarchaeological information gives indicators of the state of preservation of the remains. We will present the results of work in 2011 with a focus on the integration of geophysical, field walking and geoarchaeological information at varying scales. Finally, we go on to discuss the non-technical issues of integrating work and measurements across different scales, time periods, working groups and academic disciplines.

1.1 Background

In the first decade of this century, the Groningen Institute of Archaeology conducted extensive field survey work in the basin of the Raganello river in northern Calabria. GIA involvement actually began in 1991 with excavations on the Timpone della Motta (a Bronze Age to early Greek period hilltop settlement with sanctuary and necropolis) that continued until 2010. Surveys in the immediate surroundings of the Timpone were begun in 1995 but large scale systematic surveys did not start until 2000. These were initially focused on areas of the foothill zone within 5 km (the assumed ‘catchment area’) of the Timpone,
based on Central Place theory (Fig. 1). The surveys were systematic and intensive, with typical collection units of circa 50x50m, and walkers at 10m intervals. All materials were collected and studied, rather than targeting only diagnostic pottery from a particular time-period. From 2004 on, surveys were extended to two upland transects as well; no such extension into the coastal plain was possible because of the deep sedimentation. All in all, between 2000 and 2009, field teams spent a total of about six months in the field, surveying 13 km². This work yielded almost two tonnes of finds and identified 220 new archaeological sites. A synthesis of the surveys up to 2005 has been published (Attema et al. 2010) and a sites-and-finds catalogue is in preparation (van Leusen et al., in prep).

Over this decade, the project goals shifted in response to the data being gathered and the mounting problems in interpreting these without considering the post-depositional history of the sites. Initially, the surveys were aimed at ‘finding out what happens in immediately around the Timpone della Motta site’ to ‘mapping all traces of settlement and land use in all of the different landscape units of the Raganello watershed’. Finally, in our current research, we aim to understand how depositional and post-depositional processes bias our record of the surface archaeology in this type of Mediterranean landscape’. The latter goal requires dealing with scales (and phenomena) that run from sub-metre to more than 1km, and is at the core of our current project ‘Rural Life in Protohistoric Italy’². We aim to tackle the problem by considering in detail the protohistoric (Bronze Age to Archaic) rural sites we’ve identified in this landscape, since these present the greatest problems of visibility and interpretation. The long GIA landscape research program has resulted in a GIS repository of a complex set of interrelated data, combining legacy data, third-party data and data we have acquired and transcribed ourselves. This paper will focus on the technical and theoretical integration of these highly varied datasets, and the pitfalls and benefits thereof.

1.2 Datasets Combined in our GIS

As figure 2 shows, we use data of the following loose types:

- Cartographic
- Airborne
- Archaeological
- Geology

Figure 1. Part of northern Calabria showing the study area (hashed), the major rivers, the Timpone and its 5km catchment (dashed circle) and the three major landscape zones a) coastal plain b) foothills and c) mountains.

Figure 2. Data types used in the project and their relationships to each other.

2 NWO (Netherlands Foundation for Scientific Research) grant 360-61-010
• Geomorphology
• Geophysical

Figure 2 is not intended as a formal diagram of the geographical database, but it does show how some of the datasets relate to, or arise from each other. For example, the geomorphological mapping was created using original field research in combination with the LiDAR data and Geological maps available for the area. What follows is a brief description of the actual datasets that fall into each of these categories, with discussions of how we can employ them in an integrated way to explore the archaeological and methodological questions we have.

1.2.1 Cartography

Relationships with local Comune (municipalities) give us access to cadastral data. This lets us identify who owns a particular plot of land, which is vital for getting permission to work but also for conducting interviews with owners’ parents and grandparents about the history of the land-use, and any major landscaping or changes in recent memory. It also allows us to track property divisions and mergers, which can be useful when identifying post-depositional changes.

We also have the 1955 1:10,000 scaled topographic map series, which shows terrace walls, wells, masserie (farmsteads), mulattiere (mule paths) and tratture (drove roads). These historical maps are very important because they preserve information about historical land use and agricultural practice which is still visible in our geophysical datasets (for example), and is still partially preserved in the modern landscape, but is mostly missing from modern maps. This is particularly important for understanding vertical transhumance, which we believe was a vital factor in this landscape right back into the protohistoric period. In contrast, the 1990 1:10,000 scaled topographic map series allows us to examine changes that have occurred since the 1950’s. This is very important to our research into site formation processes and detection potential, but also gives us current field boundaries and access routes.

1.2.2 Airborne data

We have a digitised archive of aerial photographs that date from 1943 onwards, though much of the earlier imagery is somewhat blurred and very difficult to accurately georeference due to the absence of suitable ground control points. We also have a series of colour orthophotos that formed the basis for the 1990 topographic mapping, which are very well georeferenced and provide vital information about land-use and land cover. If we have a series of photographs covering the same area, we can use them to track changes in land use even where the field boundaries do not change on the topographic maps. For example, in the 1980’s parts of the landscape were heavily remodelled under the influence of EU subsidies for the development of Agriturismos (agricultural tourism establishments). Large areas of scrub in part of our study area were levelled off and turned into olive groves, with obvious and serious implications for any surface-based survey techniques. We also have a LiDAR dataset acquired in association with the Catholic University of Leuven in 2008 as part of the 2-year RAGALIRS project (now extended for another 2 years), which aims to investigate how remotely sensed data may be used to identify disturbed archaeological layers in images of the topsoil. The LiDAR dataset not only supplies us with a high resolution digital elevation model for examining the geomorphology of the area, it also allows the identification of agricultural terraces, lime kilns, drove roads, threshing floors and other topographic (and potentially archaeological) features.

1.2.3 Archaeological datasets

The bulk of our archaeological evidence consists of the large body of survey data built up since 1995, when the first GIA surveys began.
around the Timpone (Attema et al. 2010). As discussed above, since 2000 this has involved large systematic and intensive surveys targeted at different land units within the 220 km² of the watershed basin of the Raganello, a seasonal river in northern Calabria. The field walking pragmatically identified sites based on a relative increase in the density of surface ceramics, but off-site ceramics were recorded and analysed as well. The analysis also attempted to account for factors affecting the visibility of ceramics in the topsoil at the time of each survey, to correct for biases in the resulting calculated sherd densities. The sites identified in the systematic survey are supplemented by about 20 cave and high-mountain sites in the surrounding mountains, discovered over the last 25 years and reported to us by the local ‘Sparviere’ speleological society.

These data are supplemented by sites digitised off a 1:50,000 scale published topographical survey by conducted in the 1960’s (De Rossi et al. 1969), whose area of work overlaps partly with ours. The survey team, working in the tradition of the Topographic School in Rome, identified some 700 sites, but there are almost no pre-colonial sites amongst these, and there was a heavy bias towards the coastal plain and foothills and towards sites easily accessible from the road system. The team relied heavily on the presence of upstanding and architectonic remains, and on second-hand reports. In the late 1980’s the protohistorian Peroni (Peroni and Trucco 1994) started to examine the pre-colonial period and identified what have since been considered ‘central places’ in territorial models of the settlement system for the region. However, the subsequent GIA field walking has shown these models to drastically underestimate the number of protohistoric sites in the landscape.

1.2.4 Geology and Geomorphology

Large scale (1:100,000) geological maps of Italy are published and available to researchers, and were digitised for the study region, but we quickly realised that these would not be detailed enough to show how geomorphological processes might affect the visibility of archaeological sites. The Raganello catchment is geologically and geomorphologically diverse, with a large number of processes operating over a range of scales. To investigate this further, PhD research was conducted by Feiken (in prep) under the auspices of the Hidden Landscapes project directed by Van Leusen (NWO grant number 276-61-002). This involved designing and producing landscape classification maps at 1:25,000 and 1:10,000 scale based on the underlying geology and the geomorphological characteristics and processes operating on the area in question. These maps have been used to understand both past settlement and land use, and past and present post-depositional processes that might be affecting a particular place - for example, the inflation or deflation of a soil profile due to tillage-induced erosion on slopes. The LiDAR dataset was a very important ingredient in the creation of these maps, as were site visits and examinations of soil and sediment profiles at key locations.

The landscape classifications have in turn informed the sampling strategy of our recent work, in which we are aiming to target a variety of different types of ceramic ‘sites’ that occur in the various landscape types in a sampling-based approach. While the landscape classification is useful for making these sampling decisions, we have discovered that we need even higher resolution mapping of the soils and slope processes to understand what is happening at site-level. For example, the landscape unit of ‘undulating sloping land’ that typically develops over gravel-fan deposits is in reality divided into a series of (older) agricultural terraces and lynchets that have developed since the commencement of machine-ploughing in the 1970’s. Each terrace has active zones of accumulation and erosion that affect the results of ceramic and geophysical surveys. This year, we have begun developing a series of soil and process maps at 1:1,000 scale at representative case study sites.
1.2.5 Geophysics

The ceramic surveys are biased by one factor that is impossible to overcome with visibility or coverage corrections: the fact that only protohistoric sites that are currently being affected by ploughing (or have been so affected in recent years) will be visible at the surface. Since 2005, geophysical survey has been experimented with as a means to overcome that problem at least partially. We intend to develop geophysical protocols to assist with the identification of ceramic sites in conditions of poor visibility (for example, only low-density hand-made impasto pottery is present, or the site is somewhat overgrown). Pilot studies were conducted in 2005/6 and 2010 (Ullrich and de Neef 2010; Kattenberg and van Leusen 2011; van Leusen et al. in prep), and from 2011 onwards geophysical survey has formed an integral part of the new Rural Life Project. In the 2011 season, geophysical surveys were conducted at very different scales. We collected Magnetic Susceptibility data over 8km transects in conjunction with soil mapping to explore the background variations in our study region, but we also conducted ‘conventional’ gradiometry over areas ranging in size from several ha to single 10m grids, depending on the nature of our targets and questions. We also work with the Geocycles research group led by Dr David Jordan at the University of Mainz to examine the sub-metre scale geophysical properties of layers and features within test pits and soil profiles to provide better models of our target anomalies. Furthermore, we try to link our interpretations across these scales, allowing them to inform each other; the geophysical interpretations (and survey strategies) rely heavily on the geological, soil, and geomorphological mapping discussed above.

2. Results of the 2011 Field Season

We will now look at two of the sites we investigated in 2011 to illustrate how these various scales and types of data inform our research questions and archaeological interpretations. The case studies also shed light on on-going integration problems. In 2011, we conducted three fieldwork campaigns that combined geophysical surveys, field walking surveys and coring and test pitting for archaeological and soil science investigations. We began by studying a class of ceramic sites containing remains of both handmade cooking vessels and large storage jars (dolia), within the landscape class undulating sloping land. This meant our work focused on an area called the Contrada Damale, to the north of the modern village of Francavilla Marittima, in the foothills of the Pollino. This area of USL lies on gravelly fan-delta deposits and conglomerates below a large limestone bluff called Serra del Gufo.

2.1 Pizzo Grande

In previous research, it was noticed that ceramic sites often occur at the base of agricultural terraces and lynchets. A model of the anthropogenic influence on the soil profile under mechanised ploughing was developed based on terrace models proposed by Krahtopolou and Frederick (2000) (Fig. 3). Our assumption is that any protohistoric habitation layers will have become buried under colluvium through natural slope processes. Traditional tillage, starting in protohistory and intensifying especially during the Hellenistic period when terracing probably began, would...
already have affected this archaeological record to some (unknown) degree. However, mechanised ploughing in the last 40 years has drastically changed slope profiles, and has led to the development of lynchets along property divisions or areas of different cropping regimes. Protohistoric sites are therefore regularly exposed at the back of the terrace (and therefore at the base of the terrace above) while they are buried too deeply for detection at the surface in the inflation zone at the front of the terrace.

The area of Pizzo Grande had been targeted for geophysical pilot studies in 2010 due to the presence of several small ceramic surface sites and the open, gently sloping nature of the site – aside, that is, from the lynchets that have developed, according to the landowner, since the 1970’s. During this pilot survey, a dense concentration of fragments of storage jars was noted along the base of one of these. The 2011 work at this site aimed to test our model for the formation of the terraces and deposits as outlined above. The geophysical surveys in 2010 (Fig. 4) had not shown any substantial buried remains, but there were some anomalies that might relate to ephemeral archaeological features, like small pits or concentrations of cultural material. A test trench cut across the lynchet bank proved our model correct - the protohistoric finds layer continued under the terrace and into the upper field where it was well protected by deep overburden (that also probably masked any geophysical signals the layer might otherwise produce). Undetected in any of the geophysical surveys (including additional ones in 2011 that were able to traverse the bank) was a limestone dry-stone wall within the bank. Its presence might be one reason that a property boundary and, later, a pronounced lynchet developed here, but the trench was too small to identify any clear orientation for the structure. In association with the wall were found protohistoric ceramics including an almost intact geometrically decorated terracotta horse, which dates to the start of the Iron Age.

This site was identified, investigated and interpreted based on a series of inferences and associations between the ceramic surveys, the geomorphological work and interviews with landowners about agricultural practices. The failure of the geophysics to detect the stone wall, and discrepancies between the gridded surveys conducted in 2011 and the cart-based (dGPS located) surveys in 2010 raise important issues of technical integration of multiple spatial datasets. This is exemplified by our failed efforts to locate a small pit-like anomaly that was present in both geophysical datasets, but in slightly different locations. Test pitting laid out by hand missed both anomalies by very small margins, and important lessons were learned about accuracy and margins of error. There is also a non-technical integration
point relating to the way the geophysical research is embedded within the archaeological programme. The geophysicists on the project work alongside the archaeologists rather than separately, allowing close coordination and rapid feedback of results and ground truthing. This means, for example, that the reasons for non-detection (such as the limestone wall) can be investigated directly in the field, rather than dismissing the geophysical survey as not having ‘worked’.

2.2 Portieri di Cerchiara

At our second site, large-scale gradiometer survey in 2010 revealed four sub-surface rectangular structures laying either side of a strongly responsive sinuous anomaly (Fig. 5). The field had not previously been field-walked but protohistoric pottery had been noted in several locations, and was eroding from a roadside section along the field edge. Our work there in 2011 has involved further geophysical surveys, Total Station recording of all surface material of a selected area, and test pits within geophysical anomalies. Geophysical measurements and bulk and micromorphological sampling for geoarchaeological studies were also done on the sections and surfaces of these pits. Operating at all of these different scales has allowed us to build up an understanding of the site in the Late Bronze Age (which most of the structures have been dated to), but also the post-depositional processes of erosion, colluviation and farming that have led to their complex expression at the surface. For example, the rectangular structures do not have a strong association with either surface ceramics or Magnetic Susceptibility enhancements of the topsoil, but the sinuous structure does. This can be explained by assuming that the upper layers of the rectangular buildings were, after their destruction or abandonment, deliberately cleared into a linear depression (possibly an ancient drove road) that ran though the site. The fine-grained interpretation of the sediments on the site is assisted by an appreciation of its position on a geomorphological saddle connecting two different geological units, with the associated complex pattern of colluviation and erosion.

3. Integration

These two examples highlight why it is useful to integrate the various data types and scales within one GIS and conceptual framework. We are trying to answer archaeological questions that require us to link observations in small test pits to processes occurring at the landscape scale, and vice versa. But what exactly have we had to do to achieve that integration? What are the technical and
conceptual challenges? We consider that there are three main types of problem to solve:

- Technical problems inherent in combining very diverse datasets;
- Integration of working practices: e.g. vocabularies and concepts across disciplines, languages and cultures;
- Issues with our understanding of the context of legacy and third-party data - the definitions and assumptions employed, and the quality and scope of the original data.

The technical problems of combining very diverse datasets, perhaps using different scales, projections, units and so on have been well documented and addressed by many before us. Though by now routine in many landscape archaeological projects, the amount of time and the level of expertise required should not be underestimated. For example, our series of 1950’s topographic maps spans a small and easily overlooked change in the base coordinate system used in Italy. The second point is always going to be an issue in an ever-fragmenting discipline that increasingly relies on international collaborations. We simply do the best we can and should make explicit the agreed upon definitions and concepts we arrive at within our projects. The third and final point is the one we wish to expand upon here, as we feel that it is in this area that the most critical and hardest to correct mistakes can be made.

3.1 Conceptual Integration?

We will illustrate what we mean by ‘conceptual integration’ by exploring some of the problems we had in compiling the various datasets used in the Rural Life in Protohistoric Italy project. A good example is provided by the issues we encountered with the scope and accuracy of our legacy archaeological data from the 1960’s. We could not simply digitise the sites from the published plans, ignoring the conditions and biases of their generation. The manner of data collection was largely ‘topographical’, most often recording upstanding or otherwise obtrusive remains and creating a strong bias towards the Hellenistic-Roman period. It was also not truly comprehensive: the survey focused on following up on reported observations and on areas that could be reached relatively easily on a moped. The resulting dataset therefore contains both direct observations and identifications of ‘sites’ based on second-hand or anecdotal evidence, such as ‘tomb reported here during 1950’s road construction’. For these second-hand observations we have no idea of their reliability, significance or spatial accuracy. Furthermore, we only have access to the 1:50,000 scale published ‘archaeological map’, not the 1:10,000 scale working map sheets nor the original field notes. This inevitably leads to problems associated with cartographic generalisation, and despite careful georeferencing means that the digitised ‘locations’ of some sites have been shown to be more than 100m off in reality. This causes problems when we try to revisit these sites to conduct our own field surveys (for example, to gain more detailed chronological or functional information). If the ‘site’ is not where the map says it should be, does that mean it has been destroyed or has it been misplaced? If remains are found in a previously blank area, what are the chances that these are in fact a mis-located site already recorded? The implications of this for spatial statistics approaches are obvious. Another problem is that where sites are given as areas rather than points, we frequently find that the size of the site has been greatly overestimated, again possibly because of the generalisations involved in scaling up from 1:10k to 1:50k. We can only responsibly use this data set if we operate with a full understanding of its limitations: it is quite clear that only sites whose presence, location and character have been confirmed in revisits should be used in any spatial modelling, for example.

We have found similar (but smaller) problems within our own field walking dataset, particularly from the early surveys of the
1990’s which were recorded using traditional paper mapping. As the project progressed, GPS mapping was adopted first for locating corners of survey units and other ground control points, and later for all field mapping purposes. Over time the accuracy of the GPS signal has improved, first with the removal of selective availability and then with the addition of WAAS (Wide Area Augmentation System) corrections, so re-locating some of the earlier and smaller sites remains problematic (Ryan and Van Leusen 2002). Site definition criteria in the field varied between team leaders and also changed as the project developed, and there are the usual problems with poor field sketches or mistakes made in data entry. We need to bear this in mind when comparing data from the earlier and later surveys, and also when considering the dataset as a whole. Even with GPS recording, ceramic scatters are only located with an absolute accuracy of between 3 and 5m (for the latest surveys); our current research goals require a greater precision. If we want to be able to make comparisons and associations between surface ceramics and geophysical anomalies that are often less than 2m in diameter, then we need to be able to locate the ceramic sites within that tolerance, or better. A thorough understanding of the ways the surveys were conducted can help with this.

Staying with the geophysical surveys, we find they often produce anomalies that, on investigation, are non-archaeological. These can relate to modern or sub-modern features in the landscape, such as old field drains, compaction due to paths and herding activities, or activities like tree removal. Two-way communication in the field between the geophysicist, survey archaeologist and local people is vital here, as is an understanding of the limitations of geophysical detection. This allows on-the-spot decisions about which anomalies need to be investigated further, and also where ‘negative’ geophysical results need more careful checking. As part of this methodological checking, we are doing a number of repeated geophysical surveys examining both seasonal variability and instrument variability, in order to separate this from the inherent spatial variability of the characteristics we measure. This ensures we are using the right scales and densities of measurement, and understand how reliably we can detect archaeological anomalies.

Underlying all of this is an understanding that however perfectly we digitise things, our GIS is not reality. At best, it represents those parts of reality that we are most interested in, but an awful lot has happened in the fields and landscapes we study. Each land-parcel has its own specific biography, and so may deviate from our generalised digital maps in some important way. We therefore have to act as anthropologists, looking at the impact of recent decision making and farming practice, such as fire breaks, pits, clearance cairns, the creation and removal of terraces and the use of controlled fires (and that of frequent uncontrolled fires as well). Understanding how such events are reflected in what we record on and below the surface is vital to understanding the results we get.

4. Conclusions

Our ultimate goal is to develop a methodology for the effective detection and mapping of small rural sites in mediterranean landscapes. Significant biases exist at all scales in the kinds of archaeological data sets that are typically collected in ‘landscape’ or ‘regional’ projects. If we want to challenge current interpretations of the rural landscape, we need to understand the biases that we have accumulated in our archaeological records and account for them in our interpretation. To do that we need to understand how the archaeology arrived at its current place in the plough soil - and thus the things that affect its detectability. The processes involved operate at multiple and varying scales so we utilise scale-appropriate techniques and data sets. Critically, an understanding of the ‘soft’ (as opposed to the technical) intricacies of data integration is
vital. You can have multiple perfectly co-located datasets, but if you can't understand what they mean and how they affect one another, then all of that technical effort is wasted.

References

Attema, P. A. J., G.-J. Burgers, and P. M. van Leusen. 2010. *Regional Pathways to Complexity: Settlement and land-use dynamics in early Italy from the Bronze Age to the Republican Period*. Amsterdam: Amsterdam University Press.


Feiken, R. *in prep.* "Geoarchaeological Investigations into Hidden Landscapes." PhD Diss. Groningen University.


van Leusen, P. M., A. Kattenberg, and K. Armstrong *in prep.* "Magnetic Susceptibility detection of small Protohistoric Sites in the Raganello Basin, Calabria, (Italy)." *Archaeological Prospection*.
Reconstructing the Ancient Cultural Landscape Around Pompeii in 2D and 3D: from Scientific Data to a Computer Animated Museum Exhibit

Sebastian Vogel
University of Tübingen, Germany

David Strebel
Visual Artist, Germany

Michael Märker
Heidelberg Academy of Sciences and Humanities, Germany

Florian Seiler
German Archaeological Institute, Germany

Abstract:
In a geoarchaeological research project the ancient cultural landscape around Pompeii is being reconstructed before the volcanic eruption of Somma-Vesuvius in AD 79. In a first phase the pre-AD 79 topography and paleo-environmental conditions were modelled using an inductively-derived predictive model based on stratigraphical data, digital elevation models and decision trees. These paleo-conditions are the basis for further detailed geoarchaeological research concerning the interactive human-environment relationship in the pre-Roman and Roman period. Within the framework of an exhibition of the State Museum for Prehistory in Halle (Germany), entitled ‘Pompeii, Nola, Herculaneum – Disasters at Mount Vesuvius’ our research project was presented in a 5 minute 3D computer animation. In this paper we illustrate the realization process of this animation. We emphasize on the conceptual design, the technical implementation and some unexpected methodological problems that arose especially in the framework of communicating scientific results in a didactic and visually appealing way to the mainly non-scientific audience of a museum exhibition.

Keywords:
3D Landscape Model, Computer Animation, Pompeii, AD 79, Museum

1. Introduction

1.1 SALVE research²

In 2006 the German Archaeological Institute and the Heidelberg Academy of Sciences and Humanities set up a geoarchaeological research project that aims at reconstructing the ancient cultural landscape around Pompeii before the eruption of Somma-Vesuvius in AD 79. This is done by investigating the natural environmental conditions and the anthropogenic influence on the landscape in Roman and pre-Roman times (http://www.salve-research.org). To characterize the paleo-environmental situation of the Sarno River plain a pre-AD 79 landscape model was generated based on the paleo-topography and paleo-landscape features. It utilized stratigraphical data from core drillings, a present-day digital elevation model (DEM) and a predictive modelling approach. Hence, for the first time a high resolution pre-AD 79 DEM was established for the entire Sarno River plain delineating the course of the paleo-river network and its floodplain as well as the location of the paleo-coastline (for more detailed information see

Vogel and Märker 2011; Vogel et al. 2011a; Vogel et al. 2011b).

In September 2010 SALVE research and the State Museum for Prehistory in Halle (Saxony-Anhalt, Germany) decided to present the latest project results in an exhibition about Pompeii that was planned for the end of 2011. In cooperation with a visual artist a 5 minute 3D computer animation was to be created to introduce the scientific approach of reconstructing the ancient landscape around Pompeii to a wider and mainly non-scientific audience. The objective of this paper is to summarize the experiences we made in the course of realization of this computer animation focusing on: (i) the concept used, (ii) the technical implementation and (iii) some unexpected difficulties that came up.

1.2 The Exhibition ‘Pompeii, Nola, Herculaneum – Disasters at Mount Vesuvius’

Between December 2011 and August 2012 the State Museum for Prehistory in Halle (Saxony-Anhalt, Germany) hosted one of the largest and most comprehensive exhibitions on Pompeii in recent years. This exhibition aimed at approaching Roman life in the shade of Mount Vesuvius volcano in a novel way: by putting the eruption of AD 79 in the context of a diachronic series of natural disasters that shaped the landscape and the lives of the people of Campania over several thousands of years from the 2nd millennium BC to the end of antiquity. The exhibition pays special attention to the people and their ability to deal with the everyday risk of natural disasters that derive from the landscape they live in and live from (State Conservation Office of Saxony-Anhalt, 2011).

2. Conceptual Design of the Animation

The iterative exchange between the scientist and the visual artist started with the creation of a storyboard defining the key elements of the animation and their broad sequential arrangement. This arrangement allowed to specify the conceptual approach by providing almost the final structure of the animation in an early stage. The main objective of this animation was not to be absolutely scientifically correct in sensu strictu but rather to give a mainly non-scientific audience an understanding of the scientific approach of reconstructing the ancient landscape around Pompeii and to shed light into the work of the SALVE research project.

In the following the main scenes of the animation are summerized:

- Aerial view of the present-day situation of the study area (Sarno River plain) and sites of interest (Pompeii etc.).

The Sarno River plain is situated south of the volcanic complex of Somma-Vesuvius. In the south and in the east the plain is flanked by branches of the Apennine Mountain range and to the west opens towards the Gulf of Naples. The plain has an approximate surface area of 210 km² and is drained by the Sarno River network (Seiler, 2008).

- Short introduction to important pre-conditions that enabled the investigation of the pre-AD 79 landscape.

During the eruption of Somma-Vesuvius in AD 79 the volcanic products were predominantly dispersed into south-eastern directions and thus blanketed the entire Sarno River plain. As the AD 79 volcanic material shows a characteristic stratigraphic sequence it serves as a chronostratigraphic marker to identify the pre-AD 79 surface (e.g. Lirer et al. 1993; Sigurdsson et al. 1985; Cioni et al. 1992).

- Presentation of the data basis of stratigraphical drillings and of the stratigraphic structure of the Sarno River plain.
A collection of more than 1,800 core drillings shed light into the stratigraphical structure of the Sarno River plain by providing detailed data on the thickness of the AD 79 volcanic deposits, the depth to the pre-AD 79 surface and the paleo-environmental character of the pre-AD 79 layer.

- Concise and understandable description of the methodological approach to reconstruct the pre-AD 79 topography and paleo-environmental conditions.

To generate a pre-AD 79 DEM for the Sarno River plain a geostatistical approach based on the stratigraphical drilling data, a high resolution present-day DEM and decision trees was applied. Subsequently, the pre-AD 79 DEM was combined with the drilling data regarding the character of the pre-AD 79 layer to further delineate important paleo-environmental features (for a detailed description of the methodology applied see Vogel and Märker, 2011).

- Presentation of the results of the reconstruction and future research.

A pre-AD 79 landscape model was generated including the pre-AD 79 topographic conditions, the natural course of the ancient fluvial network and its floodplain as well as the ancient coastline. In future geoarchaeological research this landscape model will be utilized to analyse for instance the rural settlement structure and the infrastructural organization of the plain.

- Schematic illustration of the AD 79 eruption of Somma-Vesuvius and cross fading into the present-day situation.

The animation could certainly not be concluded without mentioning the explosive eruption of Somma-Vesuvius in AD 79. However, due to the fact that the research project is rather focused on the paleo-environmental reconstruction of the landscape around Pompeii the eruption should only appear merely hinted in the animation. Furthermore, a scientifically correct visualization of the eruption would have exceeded the temporal and financial budget available.

For the animation we deliberately decided to move backwards chronologically beginning with the present-day situation of the Sarno River plain and showing the AD 79 eruption of Somma-Vesuvius at the end. For some viewers this order may cause some initial irritation because it stands in contradiction to the visual expectation, the AD 79 eruption to be the dramatic climax and the main precondition for the preservation of the ancient conditions and for the SALVE research project. However, as earlier mentioned, we rather wanted to focus on the step by step description of the scientific workflow. Following the principles of this workflow we had to move backwards starting from the present-day situation of the study area that exhibits particular characteristics which we utilized by applying specific hypotheses and up to date technologies to achieve certain research results. Finally, as a compromise for those who might expect the volcanic eruption at the beginning we arranged the animation with the same shot at the beginning and at the end. Hence, during the exhibition the animation could be shown in an infinite loop.

3. Technical Implementation of the Animation

In the course of practical implementation of the animation the main challenge for the visual artist was to use the scientific project data and present them in a visually appealing way according to the earlier mentioned objectives. For this the free open source 3D software Blender 2.5 was used (http://www.blender.org). In a first step a 3D animated landscape of the Sarno River plain was to be created. For this a raster based high resolution DEM of the study area was provided by the scientist. However, it turned out that it was not possible
to simply import the geographical raster data into the 3D software even though a variety of export file formats was tried. Consequently, an indirect import procedure had to be applied. The DEM was converted into a greyscale image that was imported into the 3D software. Then the 3D surface was generated by assigning an elevation value to each grey scale value. The first results using an 8 bit greyscale image were not convincing because 8 bit contain only \(2^8\) (256) grey scale values. This resulted in a strong reduction of the resolution of the digital elevation data and the final 3D surface. Consequently, a greyscale image was generated with a colour depth of 16 bit. It contained a total of \(2^{16}\) (65,536) grey scale values and thus could be imported into the 3D software without data loss. Since the export procedures of standard geographical information system (GIS) software only generate 8 bit grey scale images the following workflow was established. Firstly, the raster based elevation data were imported into the proprietary software Global Mapper 12 and converted into a Terragen terrain file format (.TER) (http://www.bluemarblegeo.com/global-mapper/). Secondly, the Terragen file was imported into the freeware program TerraConv and converted into a 16 bit grey scale image (http://koti.mbnet.fi/pkl/tg/TerraConv.htm).

After importing the 16 bit grey scale image into Blender 2.5 a polygon mesh plane was generated by displacement of the grey scale values with elevation values. The resulting wireframe model was then converted into a solid surface model. For the subsequent texturing of the 3D surface several data layers of different 2D top view textures were generated in GIS and converted into bitmap file format. In the 3D software these image based textures were mounted like a bed sheet on the 3D landscape model. In 3D computer graphics this process is called UV mapping (Fig. 2).

Using UV mapping the 3D model was superimposed with the following textures (Fig. 3):

- the aerial photograph to combine the physical 3D surface with real world imagery and to represent the present-day situation of the highly urbanized Sarno River plain,
- a colour gradation illustrating different elevation levels of the study area to represent either the present-day or the modelled pre-AD 79 topography,
- the location of the stratigraphical drillings,
- important present-day environmental features, e.g. the Sarno River network and
- the reconstructed paleo-conditions such as the ancient fluvial network and its
floodplain, the coastline or the ancient rural settlements.

Since the 3D software cannot deal with georeferenced data like a GIS does it was crucial for the UV mapping that all texture layers had the same spatial extent by having the identical proportion of pixels in length and width.

In the computer animation the reconstructed pre-AD 79 topography was finally visualized by superimposing the 3D model, generated from the present-day surface, with the image based texture of colour gradation of the pre-AD 79 elevation data (Fig. 3, bottom). This indirect method was used because the calculated elevation difference between the present-day and the pre-AD 79 surface is too small (approximately 5.6 m, Vogel et al. 2011) to be visible in the small scale viewer perspective used in the animation. Hence, the generation of a second pre-AD 79 surface model out of a 16 bit grey scale model could be avoided.

To further highlight to the audience in what stage of the animation the reconstructed pre-AD 79 landscape is introduced some visual elements were added:

- The vertical axis of the thickness of post-AD 79 deposits was displayed with an artificial exaggeration to clearly distinguish between the AD 79 volcanic deposits and the underlying pre-AD 79 surface (Fig. 4).
- A 3D model of a Roman *villa rustica* was generated and placed upon the pre-AD 79 surface as a reference for a Roman rural settlement (Fig. 5). For this purpose one of the most well-known excavated and the only authentically reconstructed Roman villa around Pompeii, the Villa Regina at Boscoreale, was taken as a model for the schematic replication.
- In order to illustrate the pre-AD 79 landscape to the audience the visual reconstruction of Somma-Vesuvius before the eruption of AD 79 was a prerequisite (Fig. 6). It has to be emphasized that this reconstruction of the volcano was solely done for didactical reasons and only in a visual and non-scientific manner as it was not within the scope of the research project.
As a strong contrast to the present-day aerial photograph that was shown in the first scenes of the animation visualizing the high degree of modern urbanization of the Sarno River plain a similar panoramic view of the plain was manipulated to represent the ancient rural landscape before AD 79. A cinematic visual effects shot (vfx-shot) was taken from the southern mountain range (Monte Pendolo) overlooking the Sarno River plain towards Somma-Vesuvius in the north. This spherical panorama was used for a matte painting, i.e. a digitally hand-painted scene in a photographic reality (Fig. 7). For the image manipulation Adobe Photoshop CS5 was used (http://www.adobe.com). In the 3D software the 2D matte painting was mapped on a 3D sphere and placed at the virtual camera position. A virtual camera rotation was then animated like on a tripod getting images from the surrounding area now from the artificial pre-AD 79 landscape.

Finally, the AD 79 eruption of Somma-Vesuvius was simulated by using the smoke system from physics and particles in Blender 2.5. Due to the large spatial extent of the eruption a large smoke cache needed to be created. Afterwards, the rendering of around 800 frames of smoke layers was started taking about one week on a 12 core processor machine.

In a final step the different elements and scenes of the animation were combined for the final rendering which was carried out using the Blender 2.5 internal render engine.

4. Discussion

One of the intrinsic properties of science and especially basic research is that very often scientific findings are exclusively recognized and discussed within the academic community. However, due to the exhibition about Pompeii at the State Museum for Prehistory in Halle (Germany) we were offered the unique chance to familiarize a wider non-scientific and non-academic audience with the scientific approach and the work of the SALVE research project. To meet this objective sometimes it was necessary not to be 100% scientifically precise at every time but rather to communicate the research findings in a didactic and visually appealing way.

During the realization of this animation some unexpected difficulties appeared such as to directly import the available high-resolution
geographical data into the 3D visualization software. Consequently, the above described indirect procedure had to be applied. The detailed description of the conceptual phase and the subsequent technical implementation aimed at demonstrating the entire workflow of this animation project. It required one year of close cooperation between the visual artist and the scientist to keep the balance between scientific standards and technical feasibility. This should help others to better appraise the time, technical equipment and manpower needed for the realization of a similar project.

5. Conclusions

The realization of a 5 minute 3D computer animation of the SALVE research project was described from the conceptual framework to the technical implementation. It was presented between 9th December, 2011 and 26th August, 2012 in an exhibition of the State Museum for Prehistory in Halle (Saxony-Anhalt, Germany) which was entitled ‘Pompeii, Nola, Herculaneum – Disasters at Mount Vesuvius’. The animation is also accessible online at http://www.salve-research.org (or alternatively at http://vimeo.com/37947310)

References


Using GIS to Reconstruct the Roman Centuriated Landscape in the Low Padua Plain (Italy)

Michele Matteazzi
University of Padua, Italy and Catalan Institute of Classical Archaeology, Spain

Abstract:
This paper deals with the application of Geographical Information Systems to landscape archaeological studies and, in particular, with researches that follow an archaeomorphological approach. The study outlines the analytical potential especially for studying ancient land divisions. The case study presented here, drawn from a PhD project, is specifically dealing with the contribution that such Systems can bring to the archaeomorphological study of a wide stretch of the alluvial plain extended to the south of the city of Padua, with the Venice Lagoon to the east and the Euganei Hills to the west (Fig. 1): in this area, the analysis of landscape features highlights the traces of ancient territorial structures organized by orthogonal axes. We think they could be recognised as land divisions carried out during Roman times.

Keywords:
Landscape Archaeology, Archaeomorphology, GIS, Centuriation, Padua Plain

1. Introduction

When Romans took the control of the Venetia in the II\textsuperscript{nd} century BC, the plain to the south of Padua was administratively divided among the ancient Venetic oppida of Patavium (Padua), Ateste (Este) and Atria (Adria) - see Fig. 2. In the I\textsuperscript{st} century BC these centres gradually increased their importance, first becoming coloniae Latiae (89 BC) and later, with Julius Caesar, being elevated to the rank of municipia (49 BC).

From this time and, in particular, from the subsequent Augustan age onwards, the archaeological data begin recording the existence of a population distributed over the greater part of the plain and which lasts until the entire II\textsuperscript{nd} century AD. An evidence that would lead to think that in this moment a major programme of territorial reorganisation must have been implemented in order to exploit the land’s agrarian potential to the full.

The real existence of this intervention, already variously hypothesized since the mid-XIX\textsuperscript{th} century, was officially confirmed in the 1972 by the finding, near the village of San Pietro Viminario, of a gromatic stone bearing the cadastral indications of a land division carried out in the area during early imperial times (Lazzaro 1971-2). Despite this finding, however, all the attempts to reconstruct the morphology and extent of this land division in detail have given until now few and uncertain results (Lazzaro 1981; Pesavento Mattioli 1984; Rosada and Bressan 2008).

For trying to better understand the characteristics of the Roman intervention in the low Padua plain, it was therefore decided to implement a new study based on the principles expressed by Landscape Archaeology that, proceeding by way of an archaeomorphological approach\footnote{Archaeomorphology is a discipline closely related to Landscape Archaeology. It considers the present landscape as shaped by a series of structural components (such as roads, paths, channels, field boundaries), that are the traces of interventions carried out at different times and by different communities. These components came over time to overlap, change and erase each other, transforming the landscape in a complex palimpsest of traces that could be investigated by way of a “stratigraphic” reading, i.e. by establishing some relative chronological sequence among the traces themselves. See Palet 1997.}, took advantage of using GIS and spatial technology (Matteazzi 2012; Fig. 3).
2. Methodology

2.1 The Use of GIS

This choice was suggested by the increasing use of GIS in Landscape Archaeology studies, and particularly in those that consider an archaeomorphological approach (Palet and Orengo 2010), being their success mainly due to the high spatial and planimetric accuracy they ensure. Actually, this aspect results essential especially in those territorial studies in which the metrological basis plays a key role, as the research on the ancient land divisions where the distances among landscape components become the most important criterion for chronological definitions (Tolba and Romano 1996; Slapšak and Štanjlič 1998; Orengo and Palet 2009; Palet et al. 2011).

These tools also allow that a large number of geographically referenced sources, needed to conduct the archaeomorphological research, can be included and analysed in a single environment, permitting a high analytical potential: thanks to the multilayered and multiscale environment they provide, in fact the material can be combined in many ways and at different scales in order to achieve a more accurate data set.

It must however be outlined that these technical abilities do not bring a new methodology in the archaeomorphological study, being their advantage the fact that allow the realization of more rapid, more precise and more comprehensive analysis.

Another important aspect of GIS is its capability in presenting the results. The ability to export graphics in multiple and high quality formats allows one to obtain excellent results in presentations, illustrations and dissemination tasks in general.

2.2 The Geo-database

The analytical capability of GIS and, above all, their reliability is based on the quality of the starting data entered into the system. For this reason, at the beginning of the research it was necessary to proceed in constructing of a suitable geo-database - in this case developed with ArcGIS 10 - which included all the most important geo-referenced information about the morphology of the study area. The input data were both raster and vector data.

The initial cartography, which served as basis to georeference other cartographic
and photographic raster sources, was the digitised version of Carta Tecnica Regionale (CTR) at 1:10.000 scale and the 2006-2007 orthophotographic series at the same scale, both provided by the Cartographic Office of Veneto Region (UCRV). Besides these, we have also used various tables of the Military Geographical Institute (IGM) cartographic series at 1:25.000 scale.

Concerning the cartographic sources, geomorphological and geological maps at 1:50.000 scale and planimetrics of the excavated Roman settlements have also been incorporated. Numerous historical maps, created between XVIIth and XIXth century have therefore been included: these were all georeferenced and rectified, but because of their high degree of imperfection, have never been used in metrological analysis, although they were considered as relevant documents on landscape history.

In respect of the photographic material, very useful for the archaeomorphological study were the aerial photographs taken between 1954 and 1955 approximately at 1:33.000 scale, provided by CNR of Padua (Fig. 4): these were scanned with a resolution of less than 1m/pixel, which allowed them to be orthorectified and georeferenced, obtaining RMSE of less than 5m. The interest of these photographs mainly lies in their date of realization, cause they impress the image of the landscape before the heavy modifications it has experienced since the 1960s. All the aerial photographs available at the UCRV have been used, in order to carry out a systematic reading of the territory searching for tracks that could be connected with ancient land use.

Vector data were extracted from 168 vector maps at 1:5.000 scale, from which it has been created a series of layers, including field boundaries, hydrology, paths, roads. Modern cadastral divisions, lithological soil and land use maps and the Venice Lagoon map, as well as elements of geomorphological (eg. alluvial ridges, ancient coastlines ...) and archaeological interest (an especially created Roman sites distribution map) were other vector layers included in the geo-database.

Finally, it was included a DTM with 5 m cells also provided by UCRV. The use of this DTM has proved crucial: thanks to a definition of this sort, the microrelief (e.g. alluvial ridges and depressions) and other morphogenetic characteristics that influence the morphology of the territory could particularly be highlighted, making it easy to connect the identified traces to the natural environment, for a better understanding of origin and evolution of the traces themselves, as well as the reasons that led them to be preserved within the present-day landscape or, on the contrary, to be erased.
Using GIS to Reconstruct the Roman Centuriated Landscape in the Low Padua Plain (Italy)
Michele Matteazzi

2.3 GIS Applications

For what more closely concerns the archaeomorphological study, the work has been based on digitised restitution of the main morphologies structuring the landscape, made from the carto-photographic base that has been created (Fig. 4). Specifically, this work has been realised through the creation of a vector polyline layer to which was joined a table reporting the type of the component (road, path, field boundary, municipal limit ...), the documentary source (cartographic or photographic) from which it was restituted, its orientation and morphologic features, eventual historical data associated and a hypothesis of chronology.

Another vector polyline layer merged all the information connected to traces and anomalies detectable from aerial photos (known or otherwise) and recognisable as human intervention in territory structuring (roads, paths, field boundaries...). Both layers were later correlated with the Roman sites layer, with the aim of defining the criteria for dating the identified traces, in particular by analysing the existing relationship among traces and distribution, chronology and (when available) orientation of the known roman sites.

Exploiting the high definition of the DTM at our disposal, we have also carried out certain types of topographic analysis, including Viewshed and Least Cost Route (LCR). The latter, in particular, was calculated from cost and friction surfaces: despite the numerous applicable cost models now available in the recent archaeological literature (De Silva and Pizzolo 2001; Van Leusen 2002; Fiz and Orengo 2008; Verhagen and Jeneson 2012), nevertheless it was necessary to create a specific model that took into account the particular geomorphological features of the study area, a low plain greatly influenced by fluvial and lagoon activity. This made it possible to reconstruct the ancient routes from *Patavium* to *Ateste* and from *Patavium* to *Atria*, lying on data provided by archaeomorphological analysis, archaeological and toponymy evidence, medieval written sources and, most important, the reconstructions of ancient environmental conditions (Fig. 2).

3. Reconstructing the Centuriated Landscape

Using this methodology, it was possible to identify the traces of a wide orthogonal field system which extends over the greater part of the study area (Fig. 5). This structure respects the same orientation followed by most of the components forming the territory main hydrological network and, in particular, by the Bovolenta channel, which seems having acted as an important axis of the field system itself.

The alignment of the majority of the known Roman sites along or near the traces belonging to this orthogonal system, the fact that the wall structures of the (few) excavated

---

**Figure 5.** Archaeomorphological analysis of the study area showing locations of the centuriated grids detected in the low Padua Plain (drawing by M. Matteazzi).
Roman settlements are perfectly oriented with it, beside the clear deformations and erasures caused by the setting of medieval road network and field systems, are good clues that could testify for its ancient origin.

Metrological analysis show that almost all the identified traces respect modular distances among them, based on multiples of a common divisor corresponding to 5 Roman actus (about 177.5 m): this would allow to recognise in this orthogonal field system an example of the most popular typology of Roman land divisions known as centuriation, in this case modulated on base units (centuriae) corresponding to rectangles of 15x20 actus (532x710m). According with the information provided by the gromatic stone of San Pietro Viminario, we can also try to recognise the main E-W axis of this centuriation (the decumanus maximus) in a road currently passing through the village of Cartura.

The grid we could reconstruct includes the entire Roman urban area of Patavium, suggesting a direct connection between this intervention of centuriation and the ancient town (Fig. 5). Such a relationship is also suggested by a straight stretch of road, recently identified as a part of the route from Patavium to Vicetia (Vicenza) mentioned in ancient itineraries (Matteazzi 2008): this road would then be exit from the Roman town acting as a decumanus (Fig. 6a).

Two other roads, one leading to the medieval castrum of Bovolenta (and recognised as a part of the Roman route to Atria) and another leading to the medieval fortified village of Piove di Sacco, seem to have played a different role, since they diagonally cross the centurial grid (Fig. 6a): the first one cutting each grid unit, the other one going through groups of three grid units. This fact is noteworthy because it has been argued that centuriations could be constructed from straight stretches of road, which would act as hypotenuses, or diagonal lines, of the grid, following the gromatic process known as varatio (Roth Congès 1996; Palet and Orengo 2011). These diagonal lines could cross one or more grid units, depending on the angular relationship between the road and the land division and the grid module.

In our study area we can find out another example of road acting as diagonal line of the land division. It is a long straight track highlighted by aerial photographs southeast of the village of Agna and identified as a part of the route Patavium-Atria (Fig. 6b): hypothetically, if we extend the track line to the northwest, we can see how it crosses diagonally our grid, passing through groups of four centuriae.

Concerning the true extent of this centuriated field system, we can find its northern limit in an area north of Padua, where
the traces belonging to our structure meet those belonging to the so called “centuriation of Padua North East” (Fig. 5), a still perfectly preserved example of classical centuriation with a grid moduled on base units of 20x20 actus (Brigand 2011). In this area the archaeomorphological analysis suggests that the boundary between the two centurial systems could have corresponded to an ancient alluvial ridge of Brenta river, likely during Roman times followed by the northernmost branch of this water course known in classical sources as Meduacus Maius (Fig. 2).

To the south, the traces disappear within an area where the reading of some aerial photographs had revealed, since the 1980’s, the evidence of a land division thought to be a centuriation connected to the Roman town of Atria (Fig. 5) and known in the literature as “centuriation of Adria North West” (Masiero 1999).

In this study, all the available aerial photograms (taken from the 1950’s to the present time) have been scanned, georeferenced, orthorectified and re-analysed, in order to achieve a better definition of the known tracks and try to identify any new track.

The new analysis carried out, beside the confirmation of the existence of such a centuriation, for which an unusual grid modules of 27x27 actus (958.5 x988, 5 m) have been proven, also has led to define the limits between the “centuriation of Adria” and the one we have identified in the Padua plain. These limits can be recognised, to the east, in an ancient alluvial ridge of the Po river surely active between the Bronze Age and Early Iron Age (Piovan, Mozzi and Stefani 2010); to the north, in the final stretch of what has been identified as the palaeo-bed followed in Roman times by the northernmost branch of Adige river (Mozzi et al. 2011).

Even more interesting is to note the close connection existing between these two centurial grids. As we can see in Fig. 7, the kardines of the southern “centuriation of Adria”, if extended to the northwest, diagonally intersect the northern grid, going through groups of two centuriae. This evidence seems to suggest that the centuriation we detected in the Padua plain is newer and that it was built by using the “centuriation of Adria” as an important reference point.

4. Final Remarks

In respect to the attribution of this intervention of land division it is difficult to think about a single connection with Patavium, since the territory on which the traces can be detected includes areas that in Roman times certainly belonged to the ager of Ateste (Bosio 1992). For this reason it is considered that the centuriation identified to the south of Padua could correspond to a larger pattern of structuring of the territory, in some ways comparable to examples of very large land divisions that include, in a single cadastre, territories belonging to different communities (e.g. civitates, praefecturae... - Ceraudo and

Figure 7. The relationship between the “centuriation of Adria” (in black) and the centurial grid detected in the Padua plain (in yellow): the kardines of the first one hypothetically pass through the second one, acting as diagonal lines. The two grids are clearly separated by an ancient alluvial ridge (in red) constructed by the Po river during the Bronze age (drawing by M. Matteazzi).
Ferrari 2009). This intervention would have therefore involved, at first, the whole Padua plain, and only later would be administratively divided among the various communities which were distributed over it.

The use of a particular metrology of $15 \text{ actus}$, quite common during Caesar and Augustus times both in Italy and Spain (Palet, Fiz and Orengo 2009), and the evidence provided by archaeological data testifying a population spreading from the second half of the 1st century BC, could suggest the study area has been subjected to this intense territorial structuring phase during the principate of Augustus. A centuriation carried out at this time would be consistent, to the other hand, with some other interventions promoted in the area by Augustus himself, including the re-founding of Ateste as colonia and the subsequent settlement within its countryside of many veterans of the battle of Actium, and the almost total urban restructuring of Patavium.

References


Palet, J. M. 1997. Estudi territorial del Pla de Barcelona: Estructuració i evolució del territori entre l'època iber-
Using GIS to Reconstruct the Roman Centuriated Landscape in the Low Padua Plain (Italy)  
Michele Matteazzi

romana i l'altmedieval segles II-I aC · X-XI dC. Barcelona: Institut de Cultura.


Integrating Spatial Analyses into Foraging Societies Land Use Strategies. A Case Study from the Nalón River Basin (Asturias, North of Spain)

Alejandro García
University of Cantabria, Spain

Miguel Angel Fano
University of La Rioja, Spain

Diego Garate
Museo Arqueológico de Bizkaia, Spain

Abstract:
The development of spatial technologies allows excellent documentation in archaeological site locations. Similarly, the application of spatial analyses to archaeological evidence has provided archaeologists with large, complete and diversified corpus of data. However, those data set are usually treated as a research goal and as a final result, without any previous theoretical reflections on the historical issues to be studied and which methodology is the most suitable to address those issues. At the same time, archaeological interpretations are usually based on preconceived models, without a critical integration between spatial analyses results, archaeological evidence and theoretical paradigms. The aim of this work is to discuss the need for the development of a specific methodology for the analysis of Palaeolithic sites location, as well as the importance of a regional and integrative perspective for a better understanding of forager societies mobility strategies. A case study from the Nalón basin is presented.

Keywords:
GIS, Site Location, Regional Approach, Mobility Strategies, Cantabrian Spain, Magdalenian

1. Theoretical Framework: the Locational Approach

Research on Palaeolithic hunter-gatherer societies has mainly concentrated on the archaeological deposits formed as a consequence of the daily activities of these societies. In contrast, the location and characteristics of the sites where these deposits are found: rock-shelters and caves in the case of Cantabrian Spain, have hardly been studied systematically with the application of a specific methodology. However, these places, whatever the activity that was carried out in them, are an important component of the archaeological record, as they were chosen by the hunter-gatherer societies and this choice – as a dwelling, place to process prey, a place for refuse, or a “shrine” - should not be ignored because it is a potential source of information, occasionally perhaps more significant than that provided by the archaeological record sensu stricto.

This becomes especially important if we accept the need for a regional approach, as a long-term research strategy for the study of Palaeolithic societies. If the understanding of a Palaeolithic site includes a good knowledge of its local and regional context, as the nomadic nature of these societies seems to require, it is clearly necessary to incorporate precise information about the places where archaeological deposits exist. In this way, our hypotheses about the role played by the different sites within their social context acquire greater robustness. However, a study of this kind is not only necessary within the
framework of an openly materialistic approach as described above, as the choice of a certain place may be influenced by much more subtle factors, connected with the significance or symbolism of certain elements in the landscape, or with the archaeological sites themselves. For example, the potential amount of information that Palaeolithic rock art sites are capable of providing is evident.

The systematic analysis of the characteristics and location of archaeological sites requires the application of a specific methodology allowing an objective comparison between different sites, in the framework of a regional approach as described above (García Moreno and Fano 2011). This paper presents the preliminary results of our work in the Nalón River basin (North Spain) in order to illustrate the type of observations that can be made with this methodology, and their relevance within a regional approach to the study of the late Magdalenian societies (14,2-11,8 ky BP) living in this area.

2. Archaeological Framework: the Upper Palaeolithic in the Nalón Basin

The Nalón River Basin is an important area in Upper Palaeolithic studies in Northern Spain because of a research project initiated in the 1980s which involved the excavation of several caves and rock-shelters with thick occupation levels belonging to different periods (Fortea 1981). These levels indicate that the occupation of this area must have been important from the Solutrean onwards (around 20 ky BP) although deposits dated in the Early Upper Palaeolithic and even in the Mousterian have also been identified (Fig. 1).

The rock art ensembles (16 in total) are equally evidence for a significant occupation of the river basin before the Magdalenian period. In nearly all cases, the art is found in rock-shelters or the outer part of caves, where single pre-Magdalenian decoration phases have been recognised. For the moment it is not possible to date each of the ensembles more precisely (González Sainz 2000; Fortea 2000/01). The sites with the most figures are the caves of Santo Adriano, Lluera I, Torneiros and La Viña.

The cave of Peña Candamo is the only site with decoration phases that can be undoubtedly attributed to the Magdalenian, except for the paintings in Cueva Oscura de Ania that are of dubious authenticity (Gómez Tabanera et al. 1975). In contrast, a large number of rock-shelters and caves were occupied in the Magdalenian, particularly in its most recent phase. Finally, a smaller number of sites were used in the Azilian period.

Thus, the best documented period in the Nalón River Basin is the late Magdalenian, represented by occupation levels at seven sites and evidence of rock art in an eighth, Cueva de Peña Candamo, as cited above. This cave was decorated at different times, from the Early Upper Palaeolithic to the Late Magdalenian.
In the more recent phases, both new areas and previously used areas and panels were decorated. An example of the latter case is the “Wall of the Engravings”, with figures belonging to the different phases of decoration in the cave (Moure 1981; Corchón et al. 2011). The other sites with late Magdalenian levels included in this study are the Rock-shelter of La Viña (Forteza 1992), the caves of Las Caldas (Corchón 2007; Corchón et al. 2005), Oscura de Ania (Adán et al. 2002), and Sofoxó (Corchón and Hoyos 1972-73; González Sainz 1989), the Rock-Shelter of Entrefoces (González Morales 1992; González Sainz 1989), and the Caves of La Lluera I (Rodríguez Asensio 1990), and La Paloma (Hoyos et al. 1980; González Sainz 1989).

3. Methodology

The methodology applied in this study is based on the definition of a series of indicators that describe the position of each site and the area where they are situated objectively and quantitatively (García Moreno 2012). These indicators can be divided into Parameters, which are obtained by direct observation and measurements in the field, and Variables, which require the use of a Geographical Information System to be modelled and quantified.

The parameters being analysed are:

- **Classification**: cave, rock-shelter or open-air site.
- **Absolute altitude**: height above sea level.
- **Relative altitude**: height of the site above the base level in the valley where it is located.
- **Aspect**: direction that the cave entrance faces, in degrees.
- **Valley type**: Location of the site in the river basin: coastal plain, main valley or tributary valley.
- **Topography**: location in the valley bottom or on a hillside.
- **Landscape**: the type of environment around the site. For example, open, hilly or steep landscape, or gorge.

The variables calculated with the GIS are:

- **Potential insolation**: measured as the potential number of hours of sunlight that each site would receive in each month and season, and the annual mean.
- **Viewshed**: land surface visible from each site.
- **Visual presence**: relevance of the site in its environment.
- **Slope of terrain**: mean angle of slope where each site is located and the proportion of each type of terrain within a certain radius.
- **Access to terrain**: estimate of the distances that can be covered from each site, calculated according to the slope of terrain.

At the same time, other factors could affect or influence in the habitability conditions of the caves used for different purposes. These could be internal factors (presence of streams, episodes of flooding, cave size and shape) or external conditions (proximity to water-courses or hot springs, position on natural routes, outstanding geographical features, etc).

By calculating all the variables it is possible to define the characteristics of the area where the sites are situated, their habitability conditions and position, and also make a comparative study of different sites, objectively and precisely.

In this study, the parameters **Classification, Valley type, Topography** and **Landscape** were used. The variables analysed were **Slope of terrain** (measured by calculating...
the isochrones and the *Least Cost Path*) and the *Viewshed*. The analysis of the variables was restricted to a radius of 10km around each site. They were calculated with the ESRI ArcGIS software using the Digital Elevation Models (DEM) published by the Spanish Centro Nacional de Información Geográfica. Depending on the variable being calculated, two different DEM were used, one at a scale equivalent to 25 metre-squares and another equivalent to 5 metre-squares, allowing greater precision.

3.1 Slope

The Slope of Terrain was calculated from the slope in each DEM cell within the 10km radius and measured as a percentage. The cells in the model were later grouped in four categories, depending on their value, in order to define the type of relief in which each site is located:

- Slope < 5%: plain.
- Slope = 5% - 15%: hilly terrain.
- Slope = 15% - 30%: steep terrain.
- Slope > 30%: rocky terrain.

The use of these four categories means that the predominant terrain type in the surroundings of each site can be determined. At the same time, it allows the different sites to be compared in a much more precise way than by a simple calculation of the mean slope of terrain.

3.2 Access to Terrain: Isochrones and Least Cost Path

From our point of view, the objective of calculating the Access to Terrain is not to define the catchment area of each site (*sensu* Higgs and Vita-Finzi 1972), but to approach the way that the terrain could be exploited from each site and the preferential areas reached from the sites.

The isochrones from each site were calculated using the slopes of terrain, following the formula proposed by Uriarte (2005). Isotropic isochrones were calculated, considering that they were return journeys taking a time of about two hours.

The *Least Cost Path* was calculated by using the tangent of the slope, in order to simulate the increasing non-linear effect of height differences on movements (Bell and Lock 2000). As in previous cases, the calculation was limited to a 10km radius around each site and was based on the sum of the accumulated cost of crossing each cell in the DEM, expressed as the tangent of its slope.

3.3 Viewshed

The viewshed from each site was defined as the surface area of the terrain visible from a given point. The radius of 10km proved appropriate as this is considered the maximum mean distance at which an observer is capable of understanding what is seen (Gillings And Wheatley 2001). In addition, although the visibility from a point is not totally reciprocal, we understand the viewshed as a way of assessing the *Visual Presence* mentioned above.

The approach being proposed here is not without its limitations. In some cases, these limitations depend directly on the archaeological record itself, whereas in other cases they are associated with the methodology of analysis used. Regarding the archaeological record, it is worth mentioning the sample size – eight sites and none in the open-air – and the fact of supposing the different Magdalenian occupations are contemporary, when it cannot be guaranteed that they are strictly synchronic (Jochim 1991). Regarding the methodology used, it should be borne in mind that any study based on a GIS depends on the queries formulated, the variables included in the study and the algorithms used. On the other hand, the presence of dense vegetation may also affect visibility (Gillings and Wheatley 2001);
in this case, the use of predictive models allows approaching the potential distribution of forests (García Moreno 2008), but the low accuracy of such models prevents its use in visibility analyses. Therefore, the results obtained should not be regarded as definitive conclusions, but as data that allow hypotheses to be proposed that are increasingly well-founded.

4. Preliminary Results and Discussion

The locations of seven archaeological deposits, all situated in caves apart from the two rock-shelters, La Viña and Entrefoces, have been analysed. The location of Cueva de Peña Candamo, a site without a Magdalenian occupation but with rock art belonging to that period, has equally been studied.

The first element of classification is the location of the sites along the Nalón River valley. Three of them (Candamo, La Lluera and La Viña) are situated in the valley itself, while the other five are located in tributary valleys. All the sites except for La Viña and Peña Candamo are situated in the valley bottom; in some cases like Sofoxó and La Lluera they are practically at river level and one of the entrances of Las Caldas acts as a resurgence (Table 1).

Certain variability is also perceived in the type of terrain around each site, measured as a function of the slope. Three groups can be differentiated. First, the sites of Sofoxó, La Paloma and Oscura de Aniá are situated in wide, open valleys with gentle slopes. Areas with slopes steeper than 30% make up about 33% of the terrain within a 10km radius around these sites. Second, the group formed by Las Caldas, Peña Candamo, La Lluera and La Viña are in areas where slopes steeper than 30% amount to 40% of the surrounding terrain; despite being located in landscapes of gentle relief, these also include some quite large hills. Third, the single site of Entrefoces is located in a narrow gorge, where slopes greater than 30% make up some 60% of the terrain.

Regarding the visibility (Fig. 2), a clear difference is seen between sites with large viewsheds, which are those located on hillsides, like Peña Candamo and La Viña, or in open valleys, like Las Caldas, and those with a much

![Figure 2. La Viña (left) and Sofoxó (right) location.](image)

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Absolute Alt.</th>
<th>Relative Alt.</th>
<th>Classification</th>
<th>Topography</th>
<th>Landscape</th>
<th>Location</th>
<th>Slope &gt;30%</th>
<th>% Viewshed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Las Caldas</td>
<td>153</td>
<td>3</td>
<td>cave secondary</td>
<td>valley floor</td>
<td>hilly</td>
<td>riverine valley</td>
<td>38.71</td>
<td>1.44</td>
</tr>
<tr>
<td>La Paloma</td>
<td>165</td>
<td>1</td>
<td>cave secondary</td>
<td>valley floor</td>
<td>open</td>
<td>riverine valley</td>
<td>34.34</td>
<td>0.68</td>
</tr>
<tr>
<td>Oscura de Aniá</td>
<td>87</td>
<td>6</td>
<td>cave secondary</td>
<td>valley floor</td>
<td>steep</td>
<td>small gorge</td>
<td>33.58</td>
<td>0.08</td>
</tr>
<tr>
<td>Entrefoces</td>
<td>240</td>
<td>3</td>
<td>rock shelter</td>
<td>valley floor</td>
<td>gorge</td>
<td>gorge</td>
<td>39.08</td>
<td>0.11</td>
</tr>
<tr>
<td>Sofoxó</td>
<td>77</td>
<td>2</td>
<td>cave secondary</td>
<td>valley floor</td>
<td>hilly</td>
<td>river bank</td>
<td>32.91</td>
<td>0.03</td>
</tr>
<tr>
<td>La Lluera I</td>
<td>105</td>
<td>3</td>
<td>cave main</td>
<td>valley floor</td>
<td>open</td>
<td>river bank</td>
<td>39.90</td>
<td>0.81</td>
</tr>
<tr>
<td>La Viña</td>
<td>265</td>
<td>130</td>
<td>rock shelter</td>
<td>mid slope</td>
<td>hilly</td>
<td>confluence</td>
<td>42.67</td>
<td>3.19</td>
</tr>
<tr>
<td>Candamo</td>
<td>188</td>
<td>158</td>
<td>cave main</td>
<td>mid slope</td>
<td>hilly</td>
<td>riverine valley</td>
<td>39.83</td>
<td>2.07</td>
</tr>
</tbody>
</table>

Table 1. Locational characteristics of Nalon river basin archaeological sites
smaller viewshed located in valley bottoms, such as Entrefoces and Sofoxó (Fig. 3).

The analysis of the Access to Terrain, or the accumulated cost of moving along the main valley, shows that all these sites are located in areas with very low movement costs, even those situated in relatively remote tributary valleys, like Entrefoces. In the calculation of isochrones, two different patterns can be observed. On one hand, no other sites are located within two hours’ walking distance from the sites of Peña Candamo, La Viña and Entrefoces, which is the furthest inland and most remote site. On the other hand, all the other sites are located in the middle valley of the Nalón. These sites are more closely inter-related as they all have at least one other site within their isochrone.

In short, the late Magdalenian sites being studied exhibit significant similarities and differences regarding their location. Six of them are situated in relatively similar locations: valley bottoms, with a very low relative altitude (< 7m) allowing more direct access to their surrounding areas and in general with limited visibility from the sites. However, this varies depending on the characteristics of the surroundings of the sites: e.g. Las Caldas vs Sofoxó. The other two sites, Peña Candamo and La Viña, are situated in very different positions. They are located on hillsides, in prominent positions within the landscape, with relative altitudes of over 100m, and viewsheds that are in general much larger than those of the sites in the valley bottom. At the same time, they appear to be exclusive sites, as no other sites are found within two hours’ walking time from them. A no less important point is the presence of significant rock art ensembles in both La Viña and Peña Candamo (as also in the case of La Lluera).

These similarities and differences must have been significant in connection with the role played by each site in its social context. As a hypothesis, it might be proposed that a hierarchy of Magdalenian sites existed in the Nalón basin: some of them would have played a major role as centres for activities and gatherings (cf. Conkey 1980), whereas others may have played a more conventional part in connection with the types of activities carried out in and from the sites. A similar pattern can be observed in the neighbouring region of the Asón river basin, where El Mirón cave (a large, hillside located, settlement) is related to a set of smaller, logistical-like sites, located on valley bottoms (García Moreno and Fano 2011). In this respect, in a second phase of the study, it would be interesting to examine the location and characteristics of the sites in the Nalón basin with pre-Magdalenian rock art, as these might have formed part of the cultural context of the Magdalenian forager societies.
However, as occurs with the rest of the archaeological record, the data elicited here cannot be interpreted in an individualised way. To understand their significance it is necessary to develop tried and tested models about the role of each site in its regional context, and this objective requires the integration of all the archaeological information (Fig. 4). Only then will the type of analysis being proposed be fully applicable in terms of inference. It is therefore necessary for the "archaeological discourse" to generate this type of "social facts"; only with these can the optimum integration of information about the location and characteristics of Palaeolithic sites be achieved.

Acknowledgements

Research presented in this paper was carried out in the framework of the Project “Aplicación de nuevas tecnologías al estudio del arte paleolítico y su contexto social en el valle del Nalón (Asturias, España): 20000-12000 BP” (IP: Mª.S. Corchón), funded by the Spanish Ministry of Economy and Competitiveness, Ref. HAR2010-17916.

References


The primary objectives of this research are to develop and test an archaeological settlement pattern model designed to identify areas of high archaeological potential on the continental shelf of Southeast (SE) Alaska. Sea level history and glacial geology suggest that the archaeological record prior to 10,600 cal BP2 (9,400 RCYBP3) may be submerged on the continental shelf. To locate and test for sites older than 10,600, it is essential to extend archaeological survey to the continental shelf in areas that were either unglaciated refugia or areas that were deglaciated during the interval between 18,000 –10,600 cal BP (15,000 –9,400 RCYBP). This research facilitates the exploration and interpretation of the origins and character of early maritime adaptations along the Northwest Coast (NWC) of North America.

Archaeological sites document continuous occupation of SE Alaska following sea level rise above modern levels around 10,600 cal BP.

The few documented earlier occupations are small interior sites that include locations where people occasionally engaged in hunting bears at hibernacula. These sites, which date between 12,500 and 10,600 cal BP, demonstrate that the region was occupied at times of lower sea level (Dixon 1999; Fedje et al. 2008; Fedje and Mathews 2005). They support the possibility of submerged coastal sites located on the continental shelf where maritime subsistence resources were abundant.

High potential areas are defined based on synthesis and interpretation of archaeologically and ethnographically documented settlement patterns applied to reconstructions of the submerged landscape. Both inductive and deductive modelling methods were utilized (Verhagen and Whitley 2012); the scale of measurement is interval or ratio and both the analytic and the systemic contexts were assessed to develop the model (Kohler 1988: 35-37, Schiffer 1972). The model was produced at 500 RCY intervals from 10,790 to 18,100 cal BP (9,500 to 15,000 RCYBP).

Lost Worlds: A Predictive Model to Locate Submerged Archaeological Sites in SE Alaska, USA

Kelly R. Monteleone and E. James Dixon
University of New Mexico, USA

Andrew D. Wickert
University of Colorado, USA

Abstract:
The archaeological record of the northern Northwest Coast (NWC) extends to approximately 12,200 cal BP; however, much of the habitable area dating to before 10,600 cal BP is now submerged. Recent research indicates that large areas of Southeast Alaska and western British Columbia were glaciated from 21,000 to 17,000 cal BP, albeit with refugia (unglaciated areas able to support life) existing along the coast. By 16,000 cal BP, much of the region was deglaciated and ecologically viable for human habitation. This project develops and tests a model to identify high potential areas for the occurrence and preservation of archaeological sites on the continental shelf of Southeast Alaska. The paleolandscaes are developed from bathymetric data and paleoecological data.

Keywords:
Marine Survey, Underwater Archaeology, SE Alaska, Sea Level, Paleolandscape

2 cal BP = calibrated using Calib 6.0 (http://calib.qub.ac.uk/calib/calib.html). IntCal 09 (Reimer et al. 2009)
3 RCYBP = radiocarbon years before present
1. Study Area

The study region is part of the NWC culture area of North America that extends from the south end of Haida Gwaii (Queen Charlotte Islands), BC to the Gulf of Alaska in the north (Ames and Maschner 1999, 18), crossing the international boundary between Canada’s British Columbia (BC) and the US (Alaska). This research is centred (study area) on the Alexander Archipelago in SE Alaska, but also draws on previous work in Haida Gwaii. The model covers the entire study area including areas that are currently submerged and subaerial.

The Alexander Archipelago is a chain of over 10,000 islands and islets in SE Alaska. The islands rise to a maximum elevation of approximately 1100 m (3600 ft). The largest island is Prince of Wales Island (POWI) (Fig. 1). The Alexander Archipelago is the traditional home to Native American groups of Tlingit, Haida, and Tsimshian. The region is a coastal temperate rainforest and has a rich supply of maritime resources including shellfish, fish, marine birds, and marine mammals, along with a variety of terrestrial plants and animals (O'Clair et al. 1992). The shorelines contain numerous embayments and coves that harbour estuarine environments and biotic resources (Moss 1998, 91). The submerged nearshore landscape was previously glaciated and is characterized by fjords and submerged valleys that have been drowned by post Last Glacial Maximum (LGM) sea level rise.

There are 11 large rivers in southeast Alaska. These rivers are all on the mainland; there are no major river systems on the islands in the Alexander Archipelago (Johnson et al. 2008, 3-5). The regional absence of large sediment laden rivers indicates that no mechanism is present to route significant amounts of terrigenous sediment to the ocean. This is important because heavy sediment loads can deeply bury archaeological sites on the continental shelf, making them difficult to detect and sample.

1.1 Glaciation

Based on coral records from Barbados, the LGM occurred approximately 21,000 cal BP (Peltier and Fairbanks 2006, 3326). However, the record for the northern NWC appears to be diachronous. Evidence from Haida Gwaii suggests the maximum glacial extent was reached at approximately 19,000 cal BP (16,000 RCYBP) (Blaise et al. 1990, 292). In SE Alaska, the maximum glaciation occurred between 29,000 and 18,000 cal BP (25,000 – 15,000 RCYBP) (Clague et al. 2004, 86). This large temporal range largely results from limited research in the region (Kaufman and Manley 2004, 22; Mann and Hamilton 1995, 459).
shelf as a contiguous ice mass (Antevs 1929, 651; O’Clair et al. 1992, 21). More recent work, however, shows that the Cordilleran ice sheet merged with local valley glaciers to create discrete ice lobes that flowed westward to the sea. Between these ice lobes existed refugia, ice-free areas that continued to support life (Carrara et al. 2007, 232; Heaton and Grady 2003; Kaufman and Manley 2004, 22; Mann 1986). The presence of faunal remains dating to the LGM in both the Alexander Archipelago and Haida Gwaii document the presence of Late Wisconsin age refugia (Heaton and Grady 2003; Fedje and Mathewes 2005). Paleoenvironmental research suggests that refugia may have played a significant role in biotic colonization following deglaciation, in conjunction with subsequent northward migration of plant species following deglaciation about 16,340 cal BP (14,000 RCYBP; Ager and Rosenbaum 2007; Ager et al. 2010). This evidence indicates that these environments could have supported humans migrating along the NWC during the Late Wisconsin (Heaton and Grady 2003; Shafer et al. 2010).

1.2 Sea Level

During the LGM, eustatic (i.e. globally averaged) sea level was approximately 120 m below modern levels (Peltier and Fairbanks 2006). Local relative sea level history for SE Alaska is complicated by the glacial history. The load and westward flow of the Cordilleran Glacier on and adjacent to the region produced differential subsidence (glacial isostatic depression) and uplift (in a glacial forebulge). The complexity of the regional lithospheric structure precludes simple isostatic calculations; instead, “hinge” points between uplifting and subsiding areas are drawn empirically from local relative sea level data. This requires sea level curves to be regionalized based on these hinge areas. Baichtal and Carlson (2010) have compared the sea level record for the outer islands of the Alexander Archipelago to that of Josenhans (et al. 1997) and Hetherington and Reid (2003) for Haida Gwaii, and find that they are similar (Fig. 2). As Haida Gwaii and the Alexander Archipelago were both formed by the geologically recent accretion of new material to the North American continental margin, it is reasonable to assume that they have similar lithospheric properties that allow them to isostatically respond in a comparable way (Anderson 1991).

The sea level history for Haida Gwaii was determined using subaerial and marine cores, which record the transition from fresh to
marine sediments (Fig. 2: light grey line with diamonds; Josenhans et al. 1997). At the LGM, Haida Gwaii was raised due to a combination of isostatic uplift and eustatic sea level fall. In the early postglacial, Haida Gwaii lowered as the forebulge subsided. In the early Holocene, local relative sea level rose above the modern sea level by approximately 15 m. Finally, during the late Holocene, sea levels stabilized (Fedje et al. 2005, 24). Hetherington (et al. 2004) calculated the forebulge on Hecata Strait and Queen Charlotte Sound to be an upwarp of over 100 m, and that it lasted until 9,700 BP. The total sea level change for Haida Gwaii was 150 m from the LGM to 5000 BP (Fig. 2). There were some small (few metres) differences between the northern and southern areas of Haida Gwaii (Hetherington and Reid 2003: Fig. 2 circles).

Baichtal and Carlson (2010) have been working to develop a sea level curve for the outer islands of the Alexander Archipelago (Fig. 2: dark line with squares, Fig. 3). Carlson (2007) focuses on the elevations of shell middens to define sea level history. She identifies three terrace levels on the west side of POWI where people could have lived. The Upper Terraces are 16–21 m above modern sea level and would have been occupied during the highest sea level (9,400 to 5000 RCYBP). The middle terraces are between 8.5 and 13 m above modern sea level and would have been occupied from 5,000 to 2,000 RCYBP. Finally, the lower or modern terraces are 6–7 m above modern sea level and would have been occupied from 2,000 RCYBP to present. Baichtal and Carlson (2010) have identified 430 shell middens, of which 231 have been dated. This provides reliable dates for the part of the sea level curve that is above modern sea level. However, there are currently only two data points below modern sea level. The first is the -165 m point, which Baichtal refers to as a terrace. Carrara (et al. 2007, 235) indicates there was a minimum depression of the sea floor in Sitka Sound of 160 m. Barron et al. (2009) analysed a dated sediment core (EW0408-11JC) from the Gulf of Esquibel (Fig. 1). The core documents the change from freshwater to salt-water at 10,600 RCYBP based on a radiocarbon determination run on a marine shell. Applying the circa -600 year marine reservoir correction from Fedje et al. (1996), the date of inundation can be estimated to approximately 10,000 RCYBP or 11,325 to 11,775 cal BP. The core provides the second point below modern sea level documenting local sea level rise (Baichtal and Carlson 2010, Barron et al. 2009). When sea level was significantly lower, the Gulf of Esquibel would have been a fresh water lake (Baichtal and Carlson 2010).

### Table 1. Average calibrated 14C Dates for sites older than 9000 cal BP (Lee 2007, 32, 46).

<table>
<thead>
<tr>
<th>Region</th>
<th>Site</th>
<th>Component</th>
<th>Mean of Calibrated Age Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE Alaska</td>
<td>Ground Hog Bay 2</td>
<td>Lower</td>
<td>11,528</td>
</tr>
<tr>
<td></td>
<td>Hidden Falls</td>
<td></td>
<td>10,157</td>
</tr>
<tr>
<td></td>
<td>49 PET 408 (On Your Knees Cave)</td>
<td>Human Remains</td>
<td>11,121</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bone Tool</td>
<td>12,129</td>
</tr>
<tr>
<td></td>
<td>Chuck Lake</td>
<td>Loc 1 (midden)</td>
<td>9,204</td>
</tr>
<tr>
<td></td>
<td>K1 Cave</td>
<td></td>
<td>12,650</td>
</tr>
<tr>
<td></td>
<td>Lyell Bay</td>
<td>East</td>
<td>9,906</td>
</tr>
<tr>
<td></td>
<td></td>
<td>South</td>
<td>10,241</td>
</tr>
<tr>
<td></td>
<td>Echo Bay</td>
<td></td>
<td>9,916</td>
</tr>
<tr>
<td></td>
<td>Richardson</td>
<td></td>
<td>10,442</td>
</tr>
<tr>
<td></td>
<td>Arrow Creek 2</td>
<td></td>
<td>10,584</td>
</tr>
<tr>
<td></td>
<td>Gaadu Din Cave 1</td>
<td></td>
<td>12,683</td>
</tr>
<tr>
<td></td>
<td>Gaadu Din Cave 2</td>
<td></td>
<td>12,480</td>
</tr>
<tr>
<td></td>
<td>Werner Bay</td>
<td></td>
<td>12,481</td>
</tr>
<tr>
<td></td>
<td>Kilgii Gwaii</td>
<td></td>
<td>10,511</td>
</tr>
<tr>
<td>BC Mainland</td>
<td>Namu (ElSx1)</td>
<td></td>
<td>11,049</td>
</tr>
</tbody>
</table>

*Average of several mean calibrated age ranges

1.3 Archaeology and Ethnography

The earliest archaeological sites in the northern NWC demonstrate that the regional
The archaeological record extends to the time of LGM sea level rise. Table 1 depicts mean calibrated ages for the oldest component at sites that are older than 9,000 cal BP (Fig. 1). The mean ages of these sites were calculated following the same method as Lee (2007), but are updated to use the IntCal 09 calibration curves (Reimer et al. 2009).

The archaeological implications of sea level are important. Early researchers primarily relied on bathymetric maps to evaluate and explore the sea floor (Bailey and Flemming 2008, 2153). However, the advent of GIS modelling enables bathymetry to be more precisely correlated with other data sets, or layers. In total, there are 7,447 archaeological sites in the project’s current database. Alaska State Historic and Preservation Officer (SHPO) and BC Archaeology Branch provided the archaeological site data (AK SHPO 2009, BC Arch Branch 2009). Of these sites, 1,077 are within the projects study area. These sites are above mean low water; there are currently no underwater archaeological sites recorded within the study area, except for presumed shipwrecks.

For each of the model parameters, site-specific values were collected for the entire region. These values are slope, aspect, distance from stream, lakes, tributaries, coast and other archaeological sites, and coastal sinuosity. Each parameter utilized the modern environment. The sites within the study area also were compared to the appropriate temporally reconstruction parameters for each of the model inputs. Originally processed in ArcGIS using the “near” tool in proximity analysis, the GRASS “v.what.rast” was more efficient at calculating values from the raster layers developed for the model. The values calculated derived the weight and values used in the model via an inductive modelling approach (Kohler 1988, 37; Verhagen and Whitley 2012).

Ethnographic data were utilized to better approximate values and weights for the model utilizing a deductive modelling approach (Kohler 1988, 37; Verhagen and Whitley 2012). The purpose of using the ethnographic data in this research is twofold. First, this information provides insights for establishing a framework for environmental conditions that are important for human subsistence (Maschner 1992). In other words, similarities in site locations tend to exist due to basic human needs and resource distributions; demonstrated by the statistical analysis of known archaeological sites with the region. Secondly, ethnographic data provide insights into the seasonal round as it pertains to human land use based on the distribution of resources. NWC people used a system of ownership, or stewardship, to ensure returns and adequate storage of resources. They also responded to the uncertainties of resource abundance by incorporating an amount of flexibility and widening their resources; for example, shifting from mainly salmon to ratfish, shellfish and deer in years when salmon harvests were not as abundant (Moss 2011, 78). This information can be applied to paleolandsapes to help determine areas most likely to have been used by people in ancient times.

2. Methods

The archaeological predictive model (Fig. 4) was developed in three stages. Stage 1 is the assembly of the ArcGIS database. Stage 2 can be divided into two types of products: intermediate and final. Intermediate products are the raster data sets derived from Stage 1 data. These include slope; aspect; type of coast; coastal sinuosity; and, distance from coast, fresh water, tributary junctions, and archaeological sites. The second stage is derived using metrics that were computed in GRASS GIS (GRASS Development Team 2012) and in ArcGIS 10. The final products of stage 2 are weighted overlays that depict high potential areas for the probable occurrence of archaeological sites based on the geomorphic characteristics documented for known sites. The third stage of the modelling
process is testing. This was conducted in two ways. The first is through statistical tests including cross-validation, Kvamme’s Gain, and spatial autocorrelation. The second test was through marine geophysical surveying and bottom sampling. The surveys included multibeam sonar, side-scan sonar, sub-bottom profiles, ROV (remotely operated vehicle), and bottom sampling (Van Veen sampler). The different geophysical survey instruments provide a view of the sea floor used in a similar way to an aerial photograph (multibeam and side scan sonars), and a profile of the sediments below the sea floor (sub-bottom) similar to ground penetrating radar. The ROV provides an interactive view of the seafloor with video camera and the ability to collect small objects using a robotic arm. The model incorporates an iterative process where model values are refined to reflect testing results and can be regenerated.

2.1 Bathymetry and DEM

Bathymetry is the topography of the sea floor. Unlike land features, there are few sources from which to download compiled elevation data of the surface of sea floor. Initially, a one arc-second grid file (EPOPO1) was downloaded from the NOAA website. This file had too many data errors to be useful for this analysis. Regional data from early soundings, multibeam

Figure 4. Flow chart of three stages for developing the model.

Figure 5. Distribution of data points included in the seamless DEM (both bathymetry and land topography).

Figure 6. Comparison of spline and IDW DEM methods. A) Spline DEM. B) IDW DEM. C) IDW minus spline DEMs with land. Comparison was conducted between the IDW and a Spline version of the DEM to determine if there would be a significant difference in results between the two different methods.
sonar surveys, and other surveys in the region were downloaded as xyz files from the hydrology section of the NOAA website (NOAA 2011). These files were processed using a Python script that converted the data to point shape files. Other bathymetric data were purchased from Scientific Fishers Inc. (SciFish Inc.). These included coastal points and bathymetry points throughout the study area. The NOAA data were merged in ArcGIS 10.0 with the coastal and bathymetric data from SciFish Inc. to form a comprehensive bathymetric dataset used for modelling.

The adjacent terrestrial topography was downloaded from the USGS website (NED 2) (Gesch et al. 2002) These DEMs were merged to form a single land surface for the study area with a resolution of 25 m. The DEM was converted to points at 25 m intervals. To create the final seamless DEM including both bathymetry and land, point files were merged together to form a large (over 40 million point) data file (Fig. 5). This file was converted to a 5 m resolution raster using Inverse Weighted Distance (IWD) in ArcGIS’s 3D analyst toolbox (Fig. 6b).

The Spline DEM (ArcGIS 10 3D analysis tool) was generated at 5-metre resolution. Figure 6A is the Spline DEM using the same colour ramp as figure 6B. Figure 6C was generated by subtracting the Spline DEM from the IDW DEM. Land was added as a semi-transparent green for reference. The two DEMs are very similar other than a few smaller anomalies. The larger anomalies are outside the focus of this study. This comparative analysis demonstrates the two DEMs are analogous, and the IDW DEM is adequate for this research.

The IDW DEM and the multibeam data collected in May 2012 will be compared. Preliminary analysis indicates that the two data sets are very similar. Variables for the archaeological site predictive model were generated from the IDW DEM.

### 2.2 Water

An important factor for human settlement is the distances from streams, lakes, and tributary junctions. These fresh water-related variables constitute 35% of the final high potential model (Table 4). A suite of topologic and hydrologic analyses was used to predict where streams and lakes might have existed prior to inundation by the ocean.

Drainage paths were calculated from the IDW DEM using an improved, highly-efficient least-cost-path search (Metz et al. 2011) in GRASS GIS (GRASS Development Team 2012). Flow accumulation was calculated using a constant value for precipitation minus evapotranspiration of $4 \times 10^{-5}$ mm/s, characteristic of the region based on the TraCE-21K paleoclimate model (Liu 2009; He 2011). A 0.1 m$^3$/s discharge threshold, based on records from gaging stations in SE Alaska, were used to define the headwaters of streams. For each time period, all of the stream paths were clipped to the paleoshorelines. The shorelines were defined as the edge between the land and open water. The streams are routed through regions that were locally below sea level and could represent brackish wetlands or lakes. Tributary junctions were determined by finding the ends of stream segments that joined.

Archaeological site potential around streams was ranked from 1 (lowest) to 5 (highest) in consecutive buffer rings at 100, 500, 1000, 2000, and 3000 m. The 500 m buffer was ranked higher than the 100 m buffer based on statistical analysis of the distance from archaeological sites in the region. The median distance from sites to water sources is 418 m. Higher outlier values could be petroglyphs, pictographs, or specific resource patches or other types of sites such as caves or lithic sources that do not necessarily require water resources.

Possible lake locations were generated in a two-step process. First, depressions were
identified in GRASS GIS with the basin filling algorithm from the TerraFlow project (Arge, Toma, and Vitter 2000; Arge et al 2001; Arge et al 2003). This algorithm typically is used to fill pits caused by data errors in the DEM, but also fills enclosed depressions such as lake basins. The goal of the second step was to extract only those depressions that were lake basins. The depressions were classified as lakes based on their size and lack of compactness, the latter being given by:

\[ N = \frac{P}{2\sqrt{\pi}} \]

where \( N \) is non-compactness, \( P \) is perimetre, and \( A \) is area. \( N \) equals one for a circle, and increases as the shape becomes more oblong and complex (GRASS Development Team). Lakes tend to be circular, and small basins are more likely to be DEM artefacts, so larger and more compact regions are preferred. Lakes were defined based on a good visually-estimated match between modern lakes and those computed by the DEM, which was found for depressions that had an area >0.5 km² and a non-compactness of <5, or an area of >4 km² and a non-compactness of <2. Archaeological site potential with distance from lakes was ranked using the same buffer sizes and ranks as streams.

2.3 Coasts

Coastlines are perhaps the most important resource procurement areas for the NWC. In historic and proto-historic times, ethnographic accounts indicate that the largest native settlements were located in sheltered bays or harbours chosen for a good canoe-landing beach (de Laguna 1960, 30; Mears 1790, 109). Winter villages containing permanent houses were situated in locations where “the people prized a view of the more open water across which the canoes of their friends or their enemies might be seen approaching” (de Laguna 1960, 30). The summer fishing camps could be located further up bays at salmon streams (de Laguna 1960; Moss 1992, 7). Thus, it is not just the proximity to the coast, but the shape of coast and character of the offshore environment that is important.

The coastline was reconstructed at 500 RCY intervals based on the regional sea level curve using a contour line at 0 m from the DEM. This contour line was assumed to approximate the paleocoastline. The coastal sinuosity algorithm was then applied to the paleocoastlines. The mean distance of recorded archaeological sites from the coast is 676 m; however, 50% of the sites are within 176 m from the coast and 75% are within 550 m.

Mackie and Sumpter (2005) utilized shoreline intricacy to analyse settlement patterns on Haida Gwaii. They defined four categories of shoreline intricacy: linear, sinuous, elaborate, and intricate in increasing complexity. Their premise was that greater shoreline length near a site would allow for a more readily accessible intertidal and subtidal zone, and thus greater potential for food production and higher biodiversity (Mackie and Sumpter 2005, 350-351). They conclude that early sites (9400-9500 BP) existed on more “elaborate coasts” and late sites (2000-200 BP) were more evenly distributed between the shoreline intricacy categories.
This analysis measured coastal sinuosity using adjacent 3 km diameter circles along the reconstructed coast. The line length of the coast (lc) was divided by the diameter of the circle (d), which normalizes the values to the chosen circle of size (Fig. 7). Values were produced for the study area coastline that range from linear (1) to sinuous (8) with a mean of 3.55 (Fig. 8). This is a new algorithm developed for this study.

A proximity analysis (ArcGIS 10: Analysis Tools, Proximity, near) was run to determine the distance between the location of the recorded archaeological sites and coastal sinuosity values. The mean coastal sinuosity near archaeological sites is 2.72. A t-test was run between all of the coastal sinuosity values and coastal sinuosity values nearest to archaeological sites. The p-value for the t-test is less than 0.01, which demonstrates that the two sinuosity values are different. The average or mean coastal sinuosity in the region is 3.55, but for archaeological sites is 2.72. This suggests that people may have selected less complex coastlines for settlement and subsistence purposes. The coastal sinuosity values for the region were classified into high, medium, and low (Table 2) based on the proximity analysis from archaeological sites to coastal sinuosity. Distance from the paleoshoreline was combined with the coastal sinuosity variable to form a single ranked variable within the model. The high, medium and low ranks were then buffered in 100, 500, 1000, and 2000 m groups. (See Table 4 for the ranked values.)

### 2.4 Other Variables

From the DEM, slope and aspect were generated using ArcGIS’s 3D analyst. The best resolution for computing slope and aspect was found to be 10 m. Table 3 has the ranking of the slope and aspect used in the high potential model.

The final variable used in the model was distance between archaeological sites. Distance from known archaeological sites was buffered and ranked in 100, 500, 1000, and 5000 m. The highest weight was given to the 500 m bin. Using the concept of home range, distance from known archaeological sites is used as a variable to increase the weight for areas known to contain sites. Twenty-five percent of the sites in the study area are within 70 m of another site and 50% of the sites are within 416 m.

### Table 2. Coastal complexity ranks in relation to the location of recorded sites.

<table>
<thead>
<tr>
<th>Value</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1.5</td>
<td>Low</td>
</tr>
<tr>
<td>1.5 - 2.5</td>
<td>High</td>
</tr>
<tr>
<td>2.51 - 4.5</td>
<td>Medium</td>
</tr>
<tr>
<td>&gt; 4.5</td>
<td>Low</td>
</tr>
</tbody>
</table>

### Table 3. Example of ranking of variables (based on Maschner 1992).

<table>
<thead>
<tr>
<th>Slope</th>
<th>Degrees</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0- 2 °</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2-5 °</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5-10 °</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>10-20 °</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>20 °+</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>130°-275°</td>
<td>2 W, SW, S, SE</td>
</tr>
<tr>
<td>276°-129°</td>
<td>1 E, NE, N, NW</td>
</tr>
</tbody>
</table>

Figure 8. Coastal sinuosity of Shakan Bay area at 11,000 RCYBP.
<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Variable</th>
<th>Degrees</th>
<th>Value</th>
<th>Weights</th>
<th>Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landscape</td>
<td>Slope</td>
<td>0 - 2</td>
<td>5</td>
<td>30%</td>
<td>Calculated from DEM at 10 m resolution; values based on Maschner (1992)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.001 - 5</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.001 - 10</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.001 - 20</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.001 +</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aspect</td>
<td>130 - 275</td>
<td>2</td>
<td>5%</td>
<td>Calculated from DEM at 10 m resolution; values based on Maschner (1992)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>275.001 -129.9999</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>Distance from streams</td>
<td>100</td>
<td>4</td>
<td>20%</td>
<td>Calculated in GRASS; values based on statistical analysis of archaeological sites within region (at appropriate coastline)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3000</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distance from lakes</td>
<td>100</td>
<td>4</td>
<td>10%</td>
<td>Calculated in GRASS; values based on statistical analysis of archaeological sites within region (at appropriate coastline)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3000</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distance from coast</td>
<td>H</td>
<td>100</td>
<td>9</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>100</td>
<td>8</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>100</td>
<td>7</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distance from tributary</td>
<td>100</td>
<td>5</td>
<td>5%</td>
<td>Calculated in GRASS; values based on statistical analysis of archaeological sites within region (at appropriate coastline)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3000</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distance from archaeological sites</td>
<td>100</td>
<td>3</td>
<td>12%</td>
<td>Calculated in ArcGIS; values based on statistical analysis of distance from nearest known archaeological site (influences the model towards current land surfaces)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5000</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

*Table 4. Weights used to create final high potential model.*
2.5 High Potential Models

Each of the variables was converted into raster format, incorporating the buffers reclassified by the associated value (Table 4). In ArcGIS 10.0 spatial analyst, the weighted overlay tool was used to create high potential models for each time slice. The weights used were a combination of values derived from the statistical analysis of the known site locations in the region and the study area, the ethnographic literature, and from other models (Hamilton and Larcombe 1994; Maschner 1992). The final values of the model ranged from 0 to 4. Based on the distribution of results, moderately high potential was determined to be value 3 and high potential was determined to be value 4 (Fig. 9 a-c).

The final results were added using ArcGIS spatial analyst math tool. This produced a combined model that covered the entire study area from 10,790 to 18,100 cal BP (9,500 to 15,000 RCYBP). The results ranged from 0 to 44 (Fig. 9 b-d). Moderately high potential was determined to be above 23, and high potential was determined to be above 30. These values are slightly below the 20% threshold used in the individual models to account for the lower values further out on the continental shelf. This final model accurately incorporates the coasts, lakes, rivers, and archaeological sites in high potential areas. Both the individual temporal intervals (in 500 RCY intervals) and the combined model are being tested and utilized.

3. Results and Discussion

To test the model statistically, Kvanme’s Gain statistic was utilized (Kvanme 1988, 329; Mink, Stokes, and Pollack 2006, 235). The known 1077 sites and a 1000 randomly generated collection of points were used to calculate gain statistics (Table 5) on the modern environmental variables. All of the known archaeological site locations produced positive gain values. The random locations produced negative gain values for the moderately high potential and the combined high potential (values 3 and 4). The random locations did not produce any values for high potential sites (value 4). The combined model was not utilized in the gain test because it was created for locations that are currently underwater; this means that there was little, if any, gain using sites that are on the modern land surface. The

<table>
<thead>
<tr>
<th>Model Values</th>
<th>Data Set</th>
<th>Gain Statistic</th>
<th>Predictive Utility (gain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Known sites</td>
<td>0.5300</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Random locations</td>
<td>-3.5361</td>
<td>Negative</td>
</tr>
<tr>
<td>4</td>
<td>Known sites</td>
<td>0.9895</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Random locations</td>
<td>-</td>
<td>None</td>
</tr>
<tr>
<td>3+4</td>
<td>Known sites</td>
<td>0.5352</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Random locations</td>
<td>-3.5373</td>
<td>Negative</td>
</tr>
</tbody>
</table>

Table 5. Gain statistics for known archaeological sites and random points.
goal of the iterative process is continually to refine the model, increasing predictability and eliminating additional areas of lower potential.

Underwater predictive modelling presents several unique challenges. The greatest challenge was to develop a DEM from various sources and field-testing the DEM. Field survey is one of the primary means of testing archaeological settlement models, but is difficult and expensive in underwater applications. This model investigates depths down to 160 m (525 ft), recreational diving is safe only to 43 m (140 feet) and even then divers only stay at that depth for seven minutes. Technical and commercial divers have different limits and requirements, but training and equipment can be very expensive. In addition to the logistical constraints for recreational divers, the state of Alaska requires special training for archaeological SCUBA diving. SCUBA is not a viable option for this project, and survey must be conducted using geophysical instruments, ROVs, and bottom sampling.

Geophysical surveying is not considered equivalent to pedestrian surveys in terrestrial archaeology, and it is not equivalent to submerged archaeology either. Multibeam sonar and sub-bottom profiler provide information equivalent to similar instruments employed in terrestrial geophysics. The results of the surveys must be ground truthed and/or verified using other methods. It is essential to include the means to verify and evaluate anomalies when conducting submerged archaeology at depths below the limits of conventional diving, such as using ROV video camera investigation and sampling. Like the geophysical surveys, the ROV requires specialized technicians and is limited. The particulate matter in the water column distorts views, and the field of view is often very small, especially in deep, dark locations. These make mosaicking imagery from the ROV difficult. The ROV’s mechanical arm recovers only small objects from the sea floor, and the Van Veen samplers used collect no more than the top 25 centimetres of sediment. The costs associated with underwater archaeology are high, especially for archaeological budgets. Support vessels capable of deploying an ROV and Van Veen sampling are often too large for the slow speed required for multibeam, sidescan, and subbottom sonar surveys and often multiple vessels are required. Research conducted in remote areas further increase logistic costs.

Because marine archaeological surveying is expensive, models are needed to optimize research efforts. Models to locate submerged archaeological sites have been produced in other parts of the world (Dixon 1979; Ejstrud 2003; Evans and Keith 2011; Faught 2004; Fedje and Christensen 1999; Gaffney, Thomson and Fitch 2007; Momber 2000; Ruppe 1988). As modelling progresses, it has become increasingly apparent that they must be regionally specific and tailored to the geologic and paleoecological parameters of the region for which they are being developed. This requires generating precise regional DEMs, sophisticated understanding of local land/sea level relationships, in-depth knowledge of the regional archaeology, and accurate paleographic and paleoecological reconstructions.

The model has already proven useful for identifying specific locales for field survey. Although no conclusive archaeological sites have been located, several anomalies have been identified that require additional field-testing to determine conclusively whether or not they are archaeological sites. The iterative process is designed to increase accuracy and efficiency over time, with the ultimate goal to identify submerged sites, and extend the archaeological record seaward, into early Holocene and late Wisconsin times.

Acknowledgements

This research was supported by the National Science Foundation, Office of Polar Programs award numbers 0703980 and 1108367. The authors would also like to
thank Sealaska Heritage Institute and the two anonymous reviewers.

References


Mears, J. 1790. Voyages made in the years 1788, and 1789 from China to the north west coast of America. London: Logographic press.


Familiar Road, Unfamiliar Ground. Archaeological Predictive Modelling in Hungary

Gergely Padányi-Gulyás, Máté Stibrányi, Gábor Mesterházy
Hungarian National Museum, Hungary

Márton Deák
ELTE FFI Department of Physical Geography, Hungary

Abstract:
In Hungary there are practically no examples of archaeological predictive modelling as a cultural heritage management (CRM) tool, despite the fact that the conditions are very promising. In a country where the proportion of unidentified sites is significant, this issue is not a theoretical problem, as the chance of protecting sites of unknown location is close to zero. It was for this very reason that our workgroup at the Hungarian National Museum – National Heritage Protection Centre (HNM-NHPC) has begun research to adapt predictive modelling to Hungarian conditions, as we regard these tools as a cheap and effective way to point out places of archaeological interest in the landscape. In this paper, we would like to summarize our ongoing research in this field. We also had the opportunity to build a predictive model on LIDAR data and examined the importance of input data precision; this is also discussed here in detail.

Keywords:
Predictive Modelling, Hungary, CRM, Site-register, LIDAR

1. Introduction and Background Information

1.1 The Need for Predictive Modelling in Hungary

Site records are a significant problem for Hungarian heritage protection; one of the biggest threats is that only a relatively small number of the assumed total of archaeological sites is known. The national site register, maintained by the National Office of Cultural Heritage (NOCH), consists of more than 70,000 records, but due to overlapping and lack of spatial data in many cases, the actual known sites can only be estimated at somewhere between 25,000 and 35,000, scattered unevenly across 93,036 km² (Reményi and Stibrányi 2011; for legislation background see Wollák 2009).

As for the estimated total, many scholars have already made an extrapolation based on the number of sites on systematically surveyed areas to the whole area of the country (e.g. Jankovich-Bésán and Nagy 2004; Wollák 2004), which resulted in an estimate of 100,000–150,000 sites within Hungary. Although this estimation is only valid for sites that can be observed on the surface, this number is more or less supported by current research. Furthermore, if we take the number of actually known sites and divide it with the area of the individual counties it becomes clear that this number is uneven and varies on a large scale (Fig. 1).

The lack of archaeological data can easily be validated by examining the heritage assessment of the Hungarian section of pipeline Nabucco made by the authors. At 384 km long, it can be thought of as a transect through the county. During the research – based on systematic field survey – 339 sites were registered, of which only 77 were previously known.

The best way to handle this situation would be the continuation of regional or even country-wide multi-approach site reconnaissance campaigns. So far the first, only, and regrettably cancelled country-wide undertaking to
identify sites in Hungary, called Hungarian Archaeological Topography, operated from the 1960s to 1990 and identified about 10,000 sites on 11.7% of the area of Hungary (Laszlovszky 2003; Wollák 2009). Continuing these would be crucial, although that would only provide a solution for the long-term.

Hungarian archaeology, however, has current responsibilities due to expected delays of construction projects and more importantly the destruction of archaeological features due to little or no information regarding archaeology. We suggest that archaeological predictive modelling could be useful for such situations: if archaeologists could give reasonable forecasts regarding archaeological heritage by assessing where archaeological sites and features can be expected, this would greatly help developers to avoid affected areas in the planning phase. Our primarily goal therefore was to examine the possibilities to create accurate predictive models especially for cultural resources management (CRM) uses.

1.2 Conditions and Basis for Modelling in Hungary

The effects of natural geography on human settlements – primarily watercourses, relief features and geology – were studied from both an ethnographical (Bátky, Györfy and Viski 1941) and a geographical (Mendöl 1932) point of view in the first half of the 20th century. In terms of archaeological settlement patterns, István Méri was among the first to emphasize the significance of proximity to water and terrain features (Méri 1952), but the relationship of water and settlements is an obvious fact for archaeologists today.

Hungarian archaeological research in the past 20 years had been primarily influenced by environmental archaeology, a science closely connected to predictive modelling, but only one publication (Fekete 2008) is actually known using the method. However, in Hungary environmental reconstructions were usually also considered as predictive models. Though those are undoubtedly give limited predictions on a given coverage, we see an archaeological predictive model as a complex prediction based on several environmental factors implemented by geographical information and geostatistical analyses regarding the probable location of archaeological sites. Therefore, these environmental reconstructions must be regarded as one basis for predictive models. Nevertheless, it seems that Hungarian conditions provide good grounds for predictive modelling for at least two reasons.

A microregional cadastre established by the Geographical Institute of the Hungarian Academy of Sciences divides the country into 230 geographically homogenous small regions (Dövényi 2010). By definition, these micoregions are “homogeneous units of space with regional potential where a definitive regional ecological unity is present” (Marosi and Somogyi 1990). The size of these geographically defined areas varies between 1-1700 square kilometres, but usually about 100-500 square kilometres, which makes them fit as boundaries for the models. We must assume in advance that...
an individual model has to be made for every such microregion, but in practice probably the same model-building considerations will be valid for many environmentally similar areas.

Another great advantage is the fact that the percentage of ploughed areas is very high. According to Coordination of Information on the Environment Land Cover (CORINE) about 55% of the whole territory of Hungary is agricultural arable land. Moreover, private ownership of land is not so strong as to hinder any field survey. With an 8-days notice to the heritage authorities (also NOCH), researchers from archaeological institutions and museums are free to go anywhere they want in the area of their competence and usually do not need to ask for the landowners’ agreement beforehand. Due to the high number of ploughed areas even the Hungarian term “field survey” usually refers to plough walking. Overall, it is very easy to collect reliable archaeological input data using hand-held GPS devices with known accuracy in large areas.

2. Methodology

At this early stage our main aim was to test the adaptability of archaeological predictive modelling (Wescott and Brandon 2000; Mehrer and Wescott 2006; Verhagen 2007) for Hungarian conditions. Three pilot areas were selected, all in County Fejér: Sárrét microregion (366 km²), Sárvíz valley (624 km²) and Perkáta region (295 km²) (Fig. 2). The Sárrét microregion and the Sárvíz valley are adjacent and their models were constructed in parallel, while the model for the Perkáta region was prepared later, utilizing the earlier experiences. Lastly, a Light Detection And Ranging (LIDAR)-based predictive model was also made for the Sárvíz valley.

As our research was not supported financially, apart from the data grant for ALS survey (see below), we had to use resources and data partly from our previous surveys and projects. Due to this necessity, when selecting the study areas we had to make compromises between our scientific goals and available options. Limited resources tied our hands to some extent in other progresses as well.

2.1 Study Areas

2.1.1 Sárrét Microregion

The Sárrét microregion is at the northern part of the Transdanubian Mezőföld (meaning meadow-land) region, containing the basin of the Sárrét and the surrounding small loessy hills. For this area, only the official central archaeological database (maintained by NOCH) was available with 280 archaeological sites. A former medieval town, Fehérvár (today Székesfehérvár) was also included in the study area; therefore the effects of urban environment also became possible to examine.

2.1.2 Sárvíz Valley

This test area, also in the Mezőföld region, is adjacent to and south of the former one. The River Sárvíz runs from the aforementioned Lake Sárrét southward to the River Danube, with an approximately 100 km long course in a wide, low-lying valley with extensive wetlands.
The Sárvíz valley traverses two counties (Fejér and Tolna); the study area only contained the northern part in Fejér.

During the “Cultural landscape heritage inventory and mapping in Hungary” project (Kollányi et al. 2010) more than 200 sites were identified in the southern half of the study area during an extensive field survey by teams from the Hungarian National Museum – National Heritage Protection Centre’s (HNM-NHPC). The data was collected with handheld GPS devices, which provided the vast majority of archaeological input data. Additional archaeological data on the northern part was available from the NOCH database.

2.1.3 Perkáta Region

As with the previous regions, this area is also a part of the Mezőföld, and dominated by a network of small loess foothills and streams. Our research in the Perkáta region has focused on the surroundings of the Medieval (10th-16th centuries AD) archaeological site, Perkáta – Nyúli-dűlő, where the location of a Medieval church with more than 5500 graves were excavated in a single 3-year-long campaign between 2009 and 2011. More than 160 sites were identified here during the field campaign.

2.2 Predictive Modelling Techniques

Predictive modelling techniques can be separated into two main groups, namely the “inductive” and the “deductive” approaches (Kamermans and Wansleeben 1999), these are sometimes also referred as “correlative” and “explanatory” (Sebastian and Judge 1988) or “data driven” and “theory driven” methods (Wheatley and Gillings 2002). The main difference is the input of previously known sites: inductive models are based on the spatial distribution of these, whilst the deductive models use them only for control after model building. Therefore, in practice inductive models are more often used for archaeological predictive modelling (Verhagen 2007).

To put it simply, the main task of any predictive modelling is to offer an assessment about any given place in the study area, whether it was ideal for human occupation or not. There are different techniques that can be taken into consideration for achieving this goal (Ejstrud 2003): binary overlay, weighted binary overlay, logistic regression, Dempster-Shafer theory, Weights of Evidence etc.

As the last mentioned method is amongst the most widely applied in archaeology (e.g. Hansen 2000, Hansen et al. 2000, Diggs and Brunswig 2006) due to its GIS software implementation and inductive approach, we also decided to build our models with the archaeological adaptation of this method (Sawatzky et al. 2009).

2.3. The Weights of Evidence (WofE) Method

WofE is a multi-criteria decision-making technique (Kay and Witcher 2009), “based on the application of Bayes’ Rule of Probability with an assumption of conditional independence” (Bonham-Carter 1994; Nykänen and Salmirinne 2007). It requires two input data groups: “training points” (i.e. point representation of archaeological sites) and “evidential themes” (integer raster layers such as slope, aspect, distance to watercourses etc.). The WofE method is based on the assumption that the evidential themes are conditionally independent. For this inspection the Agterberg–Cheng Conditional Independence (CI) Test can be and was used (Agterberg and Cheng 2002).

In the first stage, a priori probability value is calculated (number of training points divided by the total area) for determining the probability of training point occurrence if they are scattered randomly. The next step is the calculation of $W^+$ and $W^-$ values for each class of each evidence themes. In case of a positive $W^+$, the number of training points is higher than expected; therefore, the class of this layer is a good positive predictor. On the other hand, negative $W^-$ values also indicate
2.4 Archaeological Data (Training Points)

The official site register of the NOCH contains spatial and metadata of archaeological research from the last 150 years. Archaeological heritage is practically the same all three regions: dense network of settlements, occupational areas and burial places from the Neolithic to the late medieval period. At that point we were not able to differentiate between types of site, however distinction was made amongst periods (see below). Accuracy and precision of these therefore vary on a very large and mostly unknown scale, mostly due the lack of GIS before the 1990s. It became quite clear that this dataset was not easily adaptable – it was also tested nonetheless – but in order to increase the model’s performance, new field surveys had to be carried out to collect accurate and precise data with handheld GPS.

The WofE method requires a point-like representation of the archaeological sites (Sawatzky et al. 2009), which raises fundamental questions: how can a single point represent an area-like site? Which point is the best for that task? Is it possible to increase the number of training points in a site? If yes, will this produce better predictions? At the very early stages of our research it was decided to test the modelling technique with different training point datasets. During the modelling of Sárvíz valley (1) every site was represented by one single point: the centroid of the polygon; (2) the training points were enriched by the number of archaeological periods found on the site, based on the assumption that if an area was ideal for settling for a longer time, it should be more pronounced; (3) the weighting was based on the size of the site. In the two latter cases the training points were divided into five categories based on major periods (all periods together, prehistory, Roman Age, Migration Period and Medieval Period), as a result, 11 training point datasets were created.

In the Sárrét microregion, the (1st) and (3rd) type were combined with the layers with and without training points from Fehérvár. The model was created, as a general model, meaning archaeological sites from all ages were used simultaneously, although afterwards Prehistoric, Roman and Medieval models were generated. Training points in the Perkáta region were generated by the (3rd) type, which eventuated four period based models.

It was also important to take into consideration that the final number of points in each training point datasets should have been between a few hundred and a maximum of a thousand (Sawatzky et al. 2009).

2.5 Environmental Data (Evidential Themes)

For describing the environment in the study areas, seven input layers were used as evidential themes during the modelling. One of the most important input layers was the Digital Elevation Model (DEM). Different kinds of DEMs were available for the study areas: Advanced Spaceborne Thermal Emission and Reflection–Global Digital Elevation Model (ASTER GDEM, ASTER 2009) for the Sárrét
microregion and the Sárvíz valley, thin plate model (Katona 2005; Neteler and Mitasova 2007) of national topographic map contour lines (1:10,000 scale) for the Perkáta region and from LIDAR data (see below) again for the Sárvíz valley. Based on these DEMs, slope, aspect and landform classification (Jennes 2006) layers were generated.

Due to the large-scale water regulations of the 19th-20th centuries, the current watercourses are not reliable data sources for historical use in the study areas (Károlyi 1973 257-262.). The utilization of DEM-based flow accumulation models was also problematic in the low-lying areas due to the accuracy and spatial resolution of DEMs. Instead of these, isotropic and anisotropic cost surfaces (Leusen 2002 6 5-7; GRASS 2011) were generated based on the water-related categories of geological mapping (Pelikán and Peregi 2005), which has the additional advantage of also containing soil types that relate to changed, ceased and dried watercourses. The contracted categories of geological mapping were also drawn into calculations at every model as a separated layer.

In the case of Sárvíz valley topsoil depth – based on the Digital Kreybig Soil Information System (Pásztor et al. 2012a; Pásztor et al. 2012b) – was also tested to localize areas identifiable with field survey. Groundwater-depth map (Ádám and Marosi and Szilárd 1959) was also available for this area (Table 1).

The WofE method requires integer rasters for evidential layers, which does not always fit pure inductive predictive models: i.e. floating value layers (e.g. slope, distance to waters etc) need to be converted into integers with 5-10 categories at maximum. This process is mainly theory driven and depends largely on the modeller himself.

2.6 LIDAR Data for Predictive Modelling

A predictive model can only be as good as the input data are. Since the DEM is the foundation of every model, it was reasonable to improve it first. At the present, LIDAR-based terrain models can produce the best terrain models even in forested areas (Crutchley 2010; Meng and Currit and Zhao 2010). The latest ALS sensors can register the full-waveform of the laser beam, and its archaeological benefits have already been researched (Doneus and Briese 2006).

ALS data in archaeology is primarily used to detect artificial features on the bare surface such as remains of buildings and roads, pits and mounds (Devereux et al. 2005; Devereux and Amable and Crow 2008; Doneus and Briese 2011). For predictive modelling, however, it
is more important that ALS data allows the construction of high resolution Digital Terrain Models (DTM) for the whole area by removing objects such as buildings and vegetation, which – in ideal circumstances – can be better by at least one order of magnitude than the best conventional elevation maps available in Hungary. Any topographic analyses such as slope, aspect, landform classification or rainfall runoff modelling will therefore produce results that are more realistic, and the predictive models generated from such layers are expected to be more functional. It was therefore possible to examine the potential of LIDAR-derived DEM usage in predictive modelling.

2.6.1 EUFAR Project ARMSRACE

Due to a data grant from EUFAR Transnational Access project called ARMSRACE (Archaeological and Relief Modelling of the Sárvíz valley for Reconstruction of Ancient Climate Events), an ALS (Airborne Laser Scanning) survey conducted by NERC (Natural Environmental Research Council) was made in August 2010 (Padányi-Gulyás and Stibrányi 2011). At the same time the LIDAR data was acquired, two 2-metre resolution hyperspectral aerial images were taken over exactly the same area with an AISA Dual sensor – one image in the visible and near infrared spectrum (VNIR, 400 – 970 nm), and one in the short-wave infrared (SWIR, 970 – 2450 nm) consisting of a total 498 bands.

2.6.2 LIDAR Data Processing

Once the unfiltered, raw point cloud was received (Fig. 3a), LIDAR processing (i.e. DTM generation by removing buildings and vegetation) had to be done before using the data for archaeological purposes. Due to disadvantageous circumstances (campaign in a month with full leaf-on conditions, higher flying altitude than expected) the point density and the vegetation penetration was below the conventional ALS datasets. Without professional LIDAR processing software the open source GRASS GIS was used, but because of the aforementioned circumstances its built-in LIDAR modules (Brovelli and Cannata and Longoni 2002; Sánchez and Brovelli 2007; Brovelli and Lucca 2009) were unable to filter out buildings and vegetation sufficiently. Therefore, it was necessary that we develop a method combining LIDAR and hyperspectral data. In the first stage, with resolution-based difference calculations on the LIDAR dataset, we were able to mask out forest and building points, but were unable to remove short and extremely dense vegetation (mainly wheat and corn, see Fig. 3b). This was done by using...
hyperspectral images: a vegetation map with relative elevations was made (Fig. 3c), and the height values were extracted from the DEM. Because of the minor errors in the hyperspectral reclassification and georeferencing, a feature-preserving denoising mesh algorithm (Sun et al. 2007) was applied, which resulted in the final DTM (Fig. 3d).

2.6.3 Topographic Analyses Based on LIDAR Data

With the LIDAR-based DTM it was possible to run topographic analyses that were impossible using the ASTER GDEM, and thereby create more precise evidential themes for a new predictive model.

Distance from water was already calculated for the Sárvíz valley from the contracted categories of geological mapping. Thanks to the detailed LIDAR-derived DTM, rainfall runoff modelling was possible by understanding more realistic flow directions and flow accumulations. An anisotropic cost-distance raster map was then created with walking time intervals.

Another advantage of the new, more detailed DTM is the opportunity to improve landform classification. Even with ASTER GDEM, some connections between archaeological sites and local hills are visible, which are presumed to be preferred for settlements. On the much more detailed new DTM, relatively small elevation anomalies were detected and clustered, strengthening this relation even more.

3. Results

With different training point datasets, different predictive models were created for every study area. Those training point datasets which sue the single point and the area-based weighting point representation of the sites are better suited for “period models” needed in academic research. Since they are optimized for a
shorter time interval, they produce better results for period-specific questions. For CRM usage, however, it was desirable to have a single model that could generally be used to avoid sites. This was reached by overlapping the area-based weighting models, using the maximum value for every period (Fig. 4). The final CRM-type models for each study area can be seen in Figs 5-8.

During the development of such CRM-type predictive models our priorities were accuracy over precision, therefore safety points like the overlapping of individual period’s models and double verification were integrated into the model building process. A two-step verification process was used to validate the models’ performance (Table 2). Initially, as

Table 2. Performance of the different models

<table>
<thead>
<tr>
<th>model</th>
<th>1st desktop verification</th>
<th>2nd field verification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>probability area %</td>
<td>control points %</td>
</tr>
<tr>
<td></td>
<td>training points</td>
<td>control points %</td>
</tr>
<tr>
<td></td>
<td>training points %</td>
<td>indicative value</td>
</tr>
<tr>
<td>S-rvÌz valley 1st version very low</td>
<td>27.19</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>35.17</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>18.98</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>18.66</td>
</tr>
<tr>
<td>S-rvÌz valley 2nd version very low</td>
<td>27.19</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>35.17</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>18.98</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>18.66</td>
</tr>
</tbody>
</table>

Table 2. Performance of the different models

Figure 7. The final DTM.

Figure 8. Landform classification.
During desktop verification, at least 72% of the training points on each model were located on 40-50% of the medium and high probability areas. The high probability categories are typically covering 5-10% of the models’ territory, showing the most affected areas, while in the very low category only a few training points and consequently few sites appear. Although the exact site territories belonged mostly to the medium and high probability categories, except from a few sites, certain “in-site” areas were marked with low or very low probability value.

4. Discussion

Compared to the others, the two models based on ASTER GDEM (Sárrét region and Sárvíz valley) produced very similar results. This strengthens our hypothesis that the DEM is one of the key components of predictive modelling. In both areas 2-3% of the training points are found in the very low, and 17-20% in the low categories (Figs 9a, 9b). This means that 19-23% of the sites could not be found in this way. The rest of the sites are allocated similarly: 57-60% in the medium and 19-21% in the high probability zone, and the same normal distribution can be observed for the ground truthing (Figs 10a, 10b).

In the Sárrét microregion, due to the lack of archaeological sites in the swampy basin, every other area with similar environmental characteristics was defined by the same low probability in the final model, which turned out
to be an error during the verification. Although involving archaeological sites from the area of the former medieval town of Fehérvár slightly improved the Kvamme’s gain values, the archaeological sites outside Fehérvár mainly belonged to non-predictive categories. Utilizing different levels of settlement hierarchy therefore distorted the modelling results (Mesterházy 2011).

In case of the ASTER-based Sárvíz valley model, there was no archaeological information about the middle area. The model worked well for avoiding sites (Fig. 11), but on the other hand, in some cases no pottery was found where the model showed highest probability. This was primarily due to the CRM-type model approach (maximized for every period), but secondly to the coarse ASTER GDEM: in case of the low-lying flat areas, the DEM was unable to differentiate slight anomalies. There was also a malfunction in one case, which turned out to be an error that occurred during the manual process of selecting water-related soil types. The modelling method therefore managed to point out such mistakes.

This picture improved significantly for the Perkáta region model, where the DEM was generated from the national topographic map contour lines. The sum of the training points found in the two lower categories is below 11%, and the graph shows a normal distribution that fully meets the needs of CRM (Fig. 9c). The same results occurred during field verification (Fig. 10c). This means that by increasing the DEM’s resolution and accuracy as well as improving the method of training point definition, the percent of sites in the two lower categories can be reduced significantly.

An interesting distribution can be seen on the graph of the LIDAR-based model (Fig. 9d). Although the values of the very low and low categories fail to show the previous results, it is evident that the model’s power is in the high category with almost 50% of the sites while keeping the Kvamme’s value above 0.6. The field verification again underlines this observation (Fig. 10d). The evaluation of this model can be best discussed by comparing it with the ASTER-based one. Such comparison showed that less than 10% of the area had higher difference than one probability category. Field verification was carried out at places where the anomaly was significant (higher than 1). These areas were mainly in the former marshland, where ASTER GDEM was unable to differentiate plains from smaller hills – which were presumed to be ideal for settling – but LIDAR could detect such microtopographic differences and the landform classification was therefore significantly improved. In these cases, the LIDAR-based model produced better results: sherds of pottery were found almost every time (Fig. 12).

The reason of this two-sided situation (worse predictive ability in the lower, but excellent in the highest category) with this last model can both be found in the nature of LIDAR and WofE. LIDAR’s greatest advantage – at least for predictive modelling – is the microtopography detection that was highly imperfect with ASTER GDEM (and sometimes even with the national topographic map contour lines), and since there really is a strong connection between archaeological sites and small hills in valleys – showed by the weight tables generated with WofE – the method regarded such features as over-important against other factors. It is possible
therefore, that the big difference of accuracy and resolution between the elevation model and the other layers distorted the results in the lower categories.

5. Conclusions

The need for a more thorough Hungarian archaeological site register, the benefits of easy field survey and the environmental factors jointly are providing a solid foundation for predictive modelling in Hungary. The borders of the Hungarian microregions are a good starting point for setting up boundaries for the single predictive models, although the Sárrét microregion shows us that sometimes microregions need to be supervised by environmental characteristics to avoid distortion in the models.

The Weight of Evidence method proved to be a useful tool for geostatistical analysis, but it is important to note that in case of floating type input layers the conversion to integer categories are not always objective. Developing, testing and verifying our models has pointed out that representing archaeological site with more than one point produces better results in the final model. During this process we enriched the training points by the size of the sites as well as the number of periods within one site.

We had assumed at an early stage that more accurate DEM implies better results in the models’ performance, which in case of Perkáta region was clearly demonstrated. With the LIDAR-based model, however, interesting results came out with worse predictive ability in the lower but better in the upper probability zones. For CRM-type predictive models the 1:10,000 national topographic map contour lines seem to be sufficient on a microregional scale, but LIDAR has unimpeachable potential for academic use.

As a result, we can confirm that predictive models could be easily adapted to Hungarian conditions, although further field surveys are required to verify our results. The definition of the key environmental variables (slope, aspect, distance to waters, geological map, landform classification) and the archaeological input data through field surveys can provide appropriate basis for modelling, and involving other layers (e.g. road networks) could considerably increase the performance of the models.

Acknowledgements

The authors would like to thank to the HNM-NHCP’s field survey team for their help and commitment and to András Zlinszky (Institute of Photogrammetry and Remote
Sensing, TU Wien) for acting as an external application advisor and remote sensing expert. We also very grateful to the two anonymous reviewers for their useful comments and especially to Roderick Salisbury for reviewing our English. The LIDAR and hyperspectral research has received funds from the European Community’s 7th Framework Programme (FP7/2008-2012). The support of the EUFAR team during the preparation, completion and aftermath of the survey is gratefully acknowledged. The NERC ARSF crew also did a wonderful job during the flight. Gergely Padányi-Gulyás would especially thank ARSF-DAN for the opportunity to visit PML and get a picture about the preprocessing. We are also grateful for Géza Király at the University of West Hungary, Sopron for testing our LIDAR data with multiple software.

References


Mathematical Models for the Determination of Archaeological Potential

Nevio Dubbini and Gabriele Gattiglia
University of Pisa, Italy

Abstract:
The Department of Archaeological Science of Pisa University is undertaking the MAPPA project in which archaeologists, geologists, mathematicians study predictive modelling tools applied to the archaeological potential of an urban area. Since the archaeological potential represents the probability that a more or less significant archaeological stratification is preserved, an analogy arose between the criteria used for attributing archaeological potential and those used for assigning importance to web pages by search engines. The key issue of the archaeological interpretation process is the identification of the relations that exist among finds, in spatial and functional terms. So different findings could strengthen or weaken the archaeological potential. Page rank algorithms are based on the same criteria: each web page attributes importance to the pages it links to and receives importance from the pages it receives a link from. This paper discusses a mathematical model of the archaeological potential based on page rank.

Keywords:
Predictive Modelling, Archaeological Potential, Page Rank, Perron-Frobenius Theorem, Mathematical Model

1. Introduction

1.1 MAPPA Project

The MAPPA (Methodologies Applied to the Archaeological Potential Predictivity, http://www.mappaproject.org) project is a research project in which archaeologists, geologists, mathematicians will study predictive modelling tools applied to the archaeological potential of an urban area, and is an answer to the problem of finding appropriate tools for making archaeological research demands coexist with present day and future needs. The Map of Archaeological Potential is the technical but above all conceptual development of common archaeological maps. All the information taken from excavations (casual or planned), ancient literary sources, archive documents, and aerial and satellite photographs are included in an archaeological map (and in the database connected to it). This is a predictive map: evaluation of the possibility that certain areas may conceal archaeological remains of which we have no news is achieved by projecting knowledge regarding neighbouring areas on them, with a degree of approximation that varies according to the quantity and quality of data available. The MAPPA project is also intended to achieve the following general objectives:

1. Enhancing the development and research of archaeology by fostering collaboration with experts from different research sectors (earth sciences and mathematics). Within this context, a predictive mathematical model applied to archaeology will have a social impact in terms of archaeological heritage protection, territorial planning and historical knowledge.

2. Creating a repeatable model that may be applied to all multi-layered urban centres in order to facilitate land use decisions regarding archaeological heritage management and protection issues. This will mitigate impact on heritage and promote the sustainability of territorial planning processes.

3. Open data: making all primary data from archaeological investigations available, proposing that after acknowledging authorship of the data, the latter shall be
made publicly available and easy to consult (Anichini and Gettiglia 2012, 54-56).

4. Training and professionally qualifying new R&D experts with a multi-disciplinary approach in order to develop a competitive system model (Anichini et al. 2011a).

1.2 Definition of Archaeological Potential

The archaeological potential of an area represents the probability that a more or less significant archaeological stratification is preserved. It is computed by analysing a series of historical, archaeological and paleoenvironmental data retrieved from various sources, with a degree of approximation that may vary according to the quantity and quality of the data provided and their spatial and contextual relationships. The archaeological potential of an area is independent of any other following intervention that is carried out, which must be regarded as a contingent risk factor. The process for defining overall urban archaeological potential consists in drawing up a series of predictive maps relative to historical periods. For this reason, besides the parameters needed for calculating the archaeological potential, parameters required for the predictive definition of the area throughout its historical periods had to be defined. This phase requires the contribution of different disciplines. The general criterion is to reconstruct stratigraphic intervals, and integrate this information with both archaeological and geomorphological data: geological maps define stratigraphic units and sedimentary bodies, geomorphological maps show relief forms and define the geomorphic processes responsible for their genesis, in addition to recent modifications. Generally speaking, each morphological unit (or morphotype) can be more or less suitable for settlements. Indeed, certain populations prefer wetlands, others favour flat areas, while others have a preference for topographic summits. Geomorphological surveying, therefore, must be based on a detailed definition of the morphological units of the current landscape and on the identification of their spatial position. Subsequently, with the help of stratigraphic data regarding the uppermost subsurface, the diachronic evolution of the forms must be characterised. A distinction must be made particularly for cases presenting: (i) continuity, over time, of geomorphic processes, yet spatial variation of forms (e.g. river processes continue to prevail yet the position of the riverbeds changes); (ii) geomorphic processes that follow on from the previous due to climatic modifications and/or crustal dynamics (e.g. marshes transformed into lagoons due to coseismic subsidence); (iii) geomorphic processes that commence or end by human intervention throughout the territory (e.g. deforestation, reclamation, etc.). In archaeological terms, the following parameters will be taken into consideration for the predictive definition of the city throughout its historical periods: typology of finds, inferred on the basis of the interpretation of the archaeological records, appropriately standardised in categories; quality and quantity of the archaeographic data; spatial and typological relations among the finds, which allow identification in probabilistic terms of the presence of further finds in areas that have not been archaeologically investigated; expert judgment; land use, including traces that are not strictly connected to constructions or settlements, such as agricultural and/or farming practices; historical data from written sources and maps.

Finally, we indentified the following overall parameters that best determine urban archaeological potential:

- **type of settlement**: the presence of settlement structures and their different typology;
- **density of settlement**: the topographic concentration of the settlement;
- **multi-layering of deposits**: the greater or lesser archaeological diachrony;
• **removable or non-removable** nature of the archaeological deposit;

• **degree of preservation of the deposit**: calculated according to the presence of anthropic and natural removals and, therefore, to the presence of documented stratigraphic gaps;

• **depth of the deposit**: this is a controversial parameter and its use alongside the other parameters mentioned above will be evaluated during the course of the project, because it appears to be strictly related to the contingency aspects of the project execution and could be confused with the calculation of archaeological risk. Instead, the parameter that will be measured is related to whether the deposit is superficial or not and to the higher or lower possibilities of it being intercepted (Anichini et al. 2011b).

### 1.3 The Data Model

The data model is developed to manage heterogeneous data, which draw the urban archaeological complexity. We need to work with both topographical (geomorphological, hydrographical, toponymic data, etc.) and urban data (archaeological stratifications, buildings, road network, hypotheses of historians and archaeologists, etc.), combining inter-site analysis and archaeological excavation GIS resources.

The archaeological data model combines raw data and interpreted data. The model goes from less synthetic data (i.e. the context level) to the more synthetic data. The centre of the model is the archaeological intervention. The data model includes also:

• filing of published data, of archive data and of data resulting from building archaeology, and data georeferencing and vectorisation in order to understand urban fabric development and the level of architectural heritage preservation;

• the collection of written, published documentary sources with the aim to locate no longer existing place names, production activities, infrastructures and topographic structures;

• the computerised acquisition of historical mapping to trace urban transformation throughout the modern and contemporary ages;

• mapping of stratigraphic gaps by recording the thick network of underground environments including cellars, garages, cisterns, basement areas, underground services, as well as the portions already subject to archaeological excavations which have led to the removal of stratigraphic deposits.

### 2. Determination of Archaeological Potential Based on Page Rank

Based on the discussions between the mathematical team and the archaeological and geological teams, an analogy arose between the criteria used for attributing archaeological potential and the criteria used for assigning importance to web pages in search engine algorithms.

The key issue of the analysis managed at § 1.2 from an abstract viewpoint is the identification of the relations that exist among the various finds, both in spatial terms (i.e. the location in space) and in functional terms (i.e. which is or could be their function). In other words, the presence of a particular find near another that has already been discovered could strengthen or weaken the probability that they will form a more complex structure, and so strengthen or weaken the archaeological potential of the area itself. This is exactly the criteria upon which page ranking algorithms are based, whereby each web page attributes importance to the web pages it points to (via a link) and, in turn, receives importance from the web pages it receives a link from.
2.1 Existing Models in Literature

Some mathematical models have already been applied to the prediction of the archaeological potential, basically divided into two groups. One of the approaches consists of a predictive model capable of generating a decision rule. The input needed to determine this rule may be, for instance, land configuration (plain/slope), the presence of nearby water sources or soil type. These features may be combined into rules such as

\[(\text{slope} \geq 10) \cap (\text{distance from source} \geq 1 \text{ km})\]

in order to predict the presence or absence of an archaeological site. Variants to this approach may include assigning ‘weights’ (i.e. more or less importance) to the different conditions, or significance statistical tests may be used (Cumming 1997; Wheatley and Gillings, 2002). Models based on these rules are very easy to implement; however, they provide on/off results – for example, the presence or absence of an archaeological site – and do not go further than simply juxtaposing a number of easy rules.

Another approach for determining archaeological potential (Wheatley and Gillings, 2002) is based on the application of linear (or logistic) regression. This approach arises from the need to reply to questions like: How can a predictor influence the model? How can continuous quantities instead of discrete quantities be predicted? Without entering into details, all linear regressions produce equations of the following type:

\[y = a + b_1 x_1 + \ldots + b_k x_k,\]

where \(y\) is the variable to be predicted (e.g. the archaeological potential), and \(x_1, \ldots, x_k\) are the inputs. In other words, by estimating coefficients \(a, b_1, \ldots, b_k\) on the basis of the available data, a \(y\) value can be found, which will be used for prediction. Several variants may also be introduced: single or multiple regressions, non-linear regressions, statistical regressions, etc. For further details on linear regression, also in relation to archaeology, please consult (Shennan, 1997; Wescott and Brandon, 2000). Although the approaches based on linear regression have the benefit of using variables to predict further variables, the model they implement is too simple and does not take into account the great complexity that must be considered when determining archaeological potential. This is so true that current models based on regression are often not preferred to those based on map algebra.

2.2 Standard Page-Rank Model

This paragraph will describe the general ideas and details of the most common mathematical models used for attributing a value of importance to web pages regardless of the value of their content and solely on the basis of the interconnections between the pages. For greater details please refer to (Langville and Meyer, 2006).

Suppose we have \(n\) web pages, numbered with integers from 1 to \(n\). We use a graph to describe the World-Wide Web, in which the nodes represent the web pages and the edges describe the links: an edge connects node \(i\) with node \(j\) if page \(i\) has a link that points to page \(j\). A directed graph may be univocally described by an adjacency matrix \(H = h_{ij}\) of size \(n \times n\) in which \(h_{ij} = 1\) if a directed edge connects node \(i\) with node \(j\), \(h_{ij} = 0\) otherwise (Fig. 1).

![Figure 1](image-url)

*Figure 1. A graphic visualization of the graph in case of 3 pages, in which page 1 points to page 2 and to page 3, page 2 points to page 1, and page 3 points to page 2.*
In the case of the graph in the figure 1 the adjacency matrix would be:

\[
H = \begin{bmatrix}
0 & 1 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix}
\]

and it represents the following system of equations:

\[
\begin{align*}
\omega_i &= \omega_2 \\
\omega_2 &= \frac{1}{2} \omega_1 + \omega_3 \\
\omega_3 &= \frac{1}{2} \omega_1
\end{align*}
\]

The criterion used to determine the importance of the web pages may be summarised as follows: a page \(i\) that points to other pages, for example \(j_1, \ldots, j_k\), distributes its importance in equal parts to pages \(j_1, \ldots, j_k\), and therefore gives \(\frac{1}{k}\) of its importance to the pages it points to.

In this model, if we call \(d_i \neq 0\) for \(i=1, \ldots, n\), and if we indicate by \(\omega_j\) the importance of page \(j\), the following holds:

\[
\omega_j = \sum_{i=1}^{n} \frac{h_{ij}}{d_i}, \quad j = 1, \ldots, n.
\]

This is simply a problem of eigenvalues and eigenvectors formulated in the following manner. If \(e = (1, 1, \ldots, 1)^T\), \(d = He\), and \(D = \text{diag}(d)\), then

\[
w^TM = w^T, \quad M = D^{-1}H,
\]

where \(w^T = (\omega_1, \ldots, \omega_n)\). This naïve formulation, however, leads to some technical problems. We list them in the following, including the adopted solutions:

1. **What happens if \(d_i = 0\) for some \(i\)?** This happens when pages do not point to anything. This is not an unusual problem, indeed, some pages (dangling nodes), such as a postscript files, may not have any links; Dangling nodes correspond to the rows having all zero entries, so a slight modification is made to the model.

The initial adjacency matrix \(H\) is replaced with \(H'\), which coincides with \(H\) everywhere apart from the zero rows, where the entries of \(H'\) are all equal to 1. It is as if we were to impartially assume that a document that in the model does not quote any other document in the web, were to quote all the existing documents in the new modified model, so it evenly distributes \(1/n\) of its importance to all of them. Matrix \(H'\) is therefore expressed as

\[
H' = H + u e^T
\]

where \(u\) is the vector with entry 1 in correspondence with the dangling nodes and zero elsewhere. Further below, matrix \(M = D'^{-1}H'\), where \(D' = \text{diag}(d')\), \(d' = He\), will be expressed as \(M\).

2. **Is there always a solution?** \(Me = e\), so 1 is an eigenvalue; consequently, \(\omega\) is any left eigenvector corresponding to the eigenvalue 1.

3. **Is the solution unique (modulo scalar multiples)?** Is the solution positive? How can it be computed? To reply to the other questions, we use the Perron-Frobenius Theorem: Let \(A\) be a \(n \times n\) matrix with non-negative entries. Then, an eigenvalue \(\lambda\) of \(A\) exists such that \(\lambda = \rho(A) \geq 0\). Also, a right eigenvalue \(x\) and left eigenvalue \(y\) correspond to \(\lambda\) with non-negative entries. Furthermore, if \(A\) is irreducible then \(\lambda\) is simple and the eigenvalues \(x\) and \(y\) have positive entries. Finally, if \(A\) has positive entries, then \(\lambda\) is the only eigenvalue of maximum module.

According to the theorem, each solution has non-negative entries. However, the non-negativity condition, on its own, does not guarantee a unique solution (modulo scalar multiples), whereas a condition of irreducibility does. It is easy to construct networks of interconnected pages that have a reducible adjacency matrix. Therefore, the model thus introduced is still not appropriate. In the case of irreducible and non-negative
matrices, other eigenvalues may exist that have the same module of the spectral radius. This leads to serious problems from an algorithmic viewpoint. In order address these problems, the matrix $M$ is replaced with the following matrix

$$A = \gamma M + (1 - \gamma) ev^T, \quad 0 < \gamma < 1,$$

where $v$ is an arbitrary vector with negative entries such that $ve^T = 1$, called personalisation vector, and $\gamma$ is a parameter, $\gamma = 0.85$ is usually chosen. Thus, matrix $A$ has positive entries. The solution exists, therefore; it is unique, and $\rho(A)$ is the only eigenvalue of module 1. From a modelling viewpoint, it is as if the importance of a page were divided into two parts: a $\gamma$ fraction is distributed on the basis of the links as in the original model, and a $1 - \gamma$ complementary fraction is distributed to all the other pages according to a criterion resulting from vector $v$. If, for example, $v = (1/n)e$, distribution shall be uniform to all web pages.

2.3 Using the Page Rank Model for Determining Archaeological Potential

In order to apply a page rank model to the determination of archaeological potential, the subsurface of the urban area is divided into $m \times n \times p$ 3-dimensional cells $(i,j,k)$, where $i,j$ define the items of a horizontal section and $k$ defines the deposit taken into consideration. The aim, therefore, is to assign the potential $x_{ijk}$ to each cell, expressed by a real non-negative number. One of the conditions we want the potential to meet is

$$x_{ijk} \leq x_{ijk'} \quad \text{if} \quad k1 \leq k2.$$

Indeed the archaeological potential of each cell should be more appropriately interpreted as the potential obtained when digging vertically from the surface down to that cell. For this reason, as the excavation goes deeper, the archaeological potential increases.

The main change to be made to the page rank model for the determination of archaeological potential regards the definition of ‘closeness’ between cells, i.e. for defining the influence that every single cell has on the other cells. In the standard page rank model, the area of influence of a web page is given by the links leaving the page itself. Building a page rank model for the archaeological potential, each node of the graph corresponds to a cell of coordinates $(i,j,k)$, and is identified by an integer $r$ such that $1 \leq r \leq mnp$. The number $r$ is determined via lexicographic ordering. Then an $N \times N$ matrix $H = (h_{rs})$ is constructed, with $N = mnp$, such that $h_{rs}$ is the part of importance that cell $r$ transfers to cell $s$. The archaeological information available for each cell is used in a twofold manner: in a relative manner, to construct the entries of $H$ controlling the transfer of importance, and in an absolute manner, giving importance to the specific cell that contains an archaeological find. In the practice, the absolute values of potential are assigned to cells that contain a specific find found in an excavation. So, after the definition of suitable parameters, there are no uncertainty in these absolute values. The estimate and the uncertainty come into play in estimation of archaeological potential for the other cells, with the matrix controlling the transfer of importance and probabilities.

The construction of the matrix $H$ is carried out on the basis of the categories used for classifying the archaeological finds. Each category characterises both the ‘geometry’ and the values of distribution of importance. For example, a cell that contains finds relating to a road may interest contiguous cells located along a line. Moreover, the normalization condition

$$\sum_{r} h_{ij} = 1$$

is maintained, which means that the entire importance of a cell is distributed: importance is neither amplified nor reduced.

Wishing to attribute a fixed value to entries $x_k$ of the archaeological potential, we may force the entry to assume the assigned value. In this case, the system $x^T H = x$ with the normalization condition $\sum x = 1$ and condition $x_k = d_k$ becomes a
linear, non-homogeneous and over-determined system, with more equations than unknowns. It must be treated, therefore, with least-squares techniques. A homogeneous approach may be maintained by renouncing matrix normalization (stochasticity). In this case, matrix $H$ is replaced by $A = DH$, $D = \text{diag}(d_1, \ldots, d_n)$ and the Perron vector of $A$ is computed, i.e. the non-negative vector $x$ such that $x^T A = cx^T$, where $\rho > 0$ is the spectral radius of $A$. Indeed, since it is no longer stochastic, matrix $A$ may not necessarily have spectral radius 1. According to this approach, by scaling the rows, we give different weights to the importance that each cell distributes to the others. A homogeneous approach may also be maintained by scaling the columns of $H$ with the assigned factor. In this case, matrix $H$ is replaced by $A = HD$. According to this approach, the importance that each cell receives is scaled (amplified or compressed). Geological information is used in a binary manner, i.e. considering or excluding cells. Alternatively, using a weight between 0 and 1 in a multiplicative manner.

The condition according to which archaeological potential increases as digging deeper into the ground may be implemented in at least two ways: by including links in the graph that connect cell $(i,j,k)$ with cell $(i,j,k+1)$; thus, the deep cells receive more importance than the surface cells; adopting the quantities $a_{ij}$ is the part of archaeological potential that cell $j$ gives to cell $i$, and such that $\sum_j a_{ij} = 1$, i.e. $A$ is stochastic. The problem to be solved, in its non-homogenous formulation, is

$$Ax = x$$

$$X_j = b_j$$

where $b_j$ is the values of archaeological potential of the cells on which we have information. Considering its homogeneous formulation, instead, the problem assumes one of the two following expressions

$$DAx = \lambda x, \quad ADx = \lambda x,$$

where $\lambda$ is the eigenvalue of maximum module of $DA$ and $AD$, respectively, whereas $D$ is the diagonal matrix whose entries give weight to the archaeological potential of the corresponding cell.

Therefore, the non-homogeneous problem is formed by an over-determined system that needs to be solved, for example, with the least-squares technique. If we replace the known values of $x_j$ it becomes a linear system in $N$ equations and $N-d$ unknowns. Instead, the homogeneous model is the solution of an eigenvalue problem.

When archaeological information is available, we proceed along two levels, as already previously described. First, we form the weight matrix $A$. Then, we assign values of importance to finds via multiplication by a diagonal matrix $D$ containing the values of importance associated to finds in corresponding cells. We could implement left or right multiplication, depending on whether we wish to load the weights on the connections that leave the cells or are received by them. Finally, the Perron eigenvalue is found.

We chose $N = 100$ to perform a simulation, and entered finds in cells 15, 37, 39, 68, with importance, respectively, of 3, 1.5, 1.7, 2. With regard to distribution of the weights, we
simulated the following cases:

- the find of cell 15 gives importance to cells 3, 4, 5, 6, 7, 8 equal to 1/6;
- the find of cell 37 gives importance to cells 45, 47, 49, 51, 53, 55, 57, 59 equal to 1/8;
- the find of cell 39 gives importance to cells 46, 48, 50, 52, 54, 56, 58, 61 equal to 1/8;
- the find of cell 68 gives importance to cells 13, 14, 15, 16, 25 equal to 1/5.

By solving the non-homogeneous and the homogeneous problems, the values of archaeological potential are obtained, as in Figure 2.

With regard to the stochastic matrix, the distribution of archaeological potential resulting from the simulation is basically distributed according to the weights assigned to each cell. Instead, the situation changes significantly if we assign importance to the finds. If we consider left-scaling – i.e. the case in which the importance of the find is ‘loaded’ on the weights received – a general increase in potential may be seen, which is higher in the areas where the finds have been discovered: the difference between the areas with greater archaeological potential and areas with lower archaeological potential is much more pronounced than in the stochastic case. With reference to the second case, archaeological potential is higher in cells 13, 14, 15, 16, 25, which receive importance in a more ‘concentrated’ manner from cell 68, which distributes a value of importance equal to 2. If we consider right-scaling – i.e. the case in which the importance of the find is ‘loaded’ on the weights sent – the iteration between the weights and the importance of the finds appears to be more significant when assigning final archaeological potential. Indeed, with regard to this case, potential is higher in cell 68, which besides having a value of importance of the find equal to 2, distributes its weight in a more ‘concentrated’ manner with respect to the other finds.

We now describe an alternative model for the determination of archaeological potential. Our aim is to provide a simple ‘basic model’, which can be used to test the goodness of the page rank-based method.

This alternative model is not built upon considerations regarding archaeological practice but consists of a ‘simple’ approximation of existing data: it is based on the principle according to which archaeological potential tends to increase by proceeding from a point where it is lower to a point where it is higher, and vice versa. We used the matlab csaps function, providing an approximation of data by means of cubic splines.

We chose \( n = 100 \) for the simulation and we included finds in cells 15, 37, 39, 68, with importance, respectively, of 3/10, 1.5/10, 1.7/10, 2/10. In order to facilitate comparison between the two different methods, the cells and the relative importance of the finds were the same used for simulation with the page rank-based models. The solution is shown in Figure 3. Since the importance of the find is the only condition that counts in this case, smoothing simply constructs a curve adapting it to the points in which the values (of importance) are known.

In Figure 4 we show the comparison the different models used, those based on the page rank model and the one based on smoothing.
The first main difference between the smoothing-based model and page-rank based model consists in the possibility to ‘distribute’ the importance of a cell to other cells, which is possible only in the page rank-based model. As already pointed out, this allows us to assign a certain ‘probability of importance’ in cells where there have been no finds. This is not possible in the smoothing-based process – nor in the models quoted in literature such as the map algebra or linear regression models. Indeed, if we observe the Figure 4, it is possible to notice how the smoothing-based model (green curve) concentrates the archaeological potential only in the areas surrounding the finds, whereas the page rank-based models concentrate the archaeological potential also around the cells which acquire importance from the cells containing the finds.

Another important difference regards the essentially relative nature of the page rank-based model, compared to the essentially absolute nature of the smoothing-based model. In the page rank-based model, in addition to the weights – which represent the portion of importance that each cell distributes to the other cells – we assigned an importance to each find by attributing a value to it which actually represented importance. This value, however, is not absolute but relative with respect to the values assigned to the other finds. Instead, in the smoothing-based model, the first problem to be addressed is the assignment of an absolute value for the potential of each cell, where known. Besides representing an additional evaluation to be carried out compared to the page rank-based model, it is also not particularly correct in theoretical and/or general terms. In fact, archaeological potential – together with characteristics that contribute to its determination (such as density of finds, rarity of finds) – is per se related to the area under examination and to the area’s size.

4. Conclusions

In conclusion, after a presentation of the MAPPA project, in which archaeologists, geologists, mathematicians aim to create a predictive modelling tool applied to the archaeological potential of an urban area, we have presented a model for estimating archaeological potential starting from already-discovered finds. We believe that existing models in literature, based mainly on map algebra and regression, are not suitable due to their extreme simplicity.

The preliminary work to be carried out in order to test these models on real data will consist in associating a ‘distribution of importance’ to each category of finds indicated by the archaeological team: the ‘distribution of importance’ will contain the information that is provided by each find about what could be discovered nearby. Finally, geological information – which will be considered in a binary manner – was not included in the simulations since implementation of this information at modelling level is immediate.
We will focus our attention on the page rank-based model because this model allows us to take into consideration a series of characteristics that follow the practical methods used by archaeologists for determining archaeological potential. The smoothing-based model, instead, which only considers finds without ‘distributing’ their importance, will be used for comparison purposes. We have already pointed out the advantages of the page-rank based models but we believe that they are much greater than initially highlighted in this phase. For example, the page rank model will be implemented considering firstly the finds in a single period and then ‘summing’ all the various periods, and bearing in mind the different interactions that finds belonging to diverse periods may have. We must also consider that the ‘distribution’ of importance may be at the level of cells but also of objects (groups of cells), i.e. at various levels of complexity.

Acknowledgements

MAPPA project has been supported by Regione Toscana and is carried out by the Department of Archaeological Sciences, Department of Earth Sciences, and the Department of Mathematics, of the University of Pisa, with the external collaboration of the Regional Directorate for Cultural and Landscape Heritage of Tuscany, of the Superintendency for Archaeological Heritage of Tuscany, of the Superintendency for Architectural, Landscape and Ethno-anthropological Heritage for the Provinces of Pisa and Livorno, of the Municipality of Pisa, of the National Institute of Geophysics and Vulcanology, of the National Aerial Photograph Archive, of the Digital Culture Laboratory – CISIAU Interdepartmental Centre of Information Services for the Humanities.

References


Calculating Accessibility

Irmela Herzog
The Rhineland Commission for Archaeological Monuments and Sites, Germany

Abstract:
Accessibility to fresh water, fertile soils, salt and other natural resources played an important role in prehistory. Accessibility is often determined in relation to target locations, for example a fresh water supply. The target locations are often points (e.g. sources) but could also be lines (e.g. streams). The simplest approach to calculate accessibility is based on the distance to the nearest target, more refined methods take several targets in the neighbourhood and their importance into account. An alternative definition of accessibility considers the ease with which a location may be reached from any other location in the area. This contribution discusses existing approaches for calculating accessibility and presents an alternative method that is closely related to kernel density estimation.

Keywords:
Accessibility, Least-Cost Paths, Kernel Density Estimation, Gravity Model, Closeness Centrality

1. Introduction

Accessibility to fresh water, fertile soils, salt and other natural resources like flint stones played an important role in prehistory. For example, accessibility to water is an important factor included in many archaeological predictive models (e.g. contributions in Kunow and Müller (2003): Dalla Bona, pp. 99–105; Ejstrud, pp. 119–34; El Kassem and Müller, pp. 135–40; Münch, pp. 171–78; Ducke pp. 185–92). Analysing accessibility is also important for modelling preferred pathways in a landscape with varying friction (Verhagen 2010).

According to Rodrigue et al. (2009, 68–69), accessibility is “the measure of the capacity of a location to be reached by, or to reach different locations”. Accessibility is measured in relation to opportunities, for example sources of water, of flint stone or other natural resources, where the location and weight indicating the importance of the opportunity form the basis for calculating accessibility. Opportunities are not necessarily points, in some cases lines or polygons might be used to model opportunities like navigable streams or subsurface strata of iron ore.

The expression “to be reached by or to reach” is reflected by two different definitions of accessibility at a given location in the landscape (Rodrigue et al. 2009, 72): (i) attractiveness of an opportunity is high, if reaching this opportunity is easy for many agents, and if the opportunity is important; (ii) emissiveness for a given location is high, if reaching (important) opportunities is easy from this location, i.e. this value measures the ease of finding opportunities (with high weight) in the neighbourhood. Generally, both variants of accessibility are defined on the basis of graphs, i.e. locations (nodes) connected by routes (edges) (e.g. Li and Shum 2001).

In addition, Rodrigue et al. (2009, 69) define contiguous accessibility, that involves measuring accessibility over a surface, i.e. a raster map is created. This paper deals with accessibility maps, i.e. contiguous accessibility. Such maps either show the emissiveness with respect to a limited set of opportunities or depict general accessibility. The term general accessibility is used for accessibility raster maps that consider each raster cell as an opportunity. Such general accessibility maps allow the identification of preferred pathways in a landscape with varying friction; but contrary to least-cost path (LCP) problems, no predefined start and end locations of the routes...
Calculating Accessibility
Irmela Herzog

need be known (Whitley and Burns 2007; Verhagen 2010).

2. Converting Cost Grids to a Graph

The basis of the accessibility maps presented in this paper are raster data storing the cost of movement between two neighbouring cells. Reconstructing the costs of movement for a given study region and period in time is not trivial (e.g. Herzog 2013). But this paper assumes that a reliable model for the costs of movement is available.

The cost of movement in the hilly study region Bergisches Land east of Cologne (Fig. 1) was derived from known Medieval trade routes recorded by Nicke (2001). The best result was achieved by LCPs avoiding wet soils and with a critical slope of 13%. Slope dependent costs were modelled by a quadratic function (Herzog 2013) and for wet soils a multiplier of 5 was used (see also Herzog, Least-cost networks, this volume Fig. 1).

The next sections describe and discuss several measures of accessibility. An overview of these measures is presented in Table 1. The

Figure 1. Hilly study region east of Cologne, Germany: (a) Overview, Cologne is marked by a lozenge. (b) The ASTER DEM is shown with a resolution of about 30 m, this DEM was supplied by NASA and METI. The altitude in the study area varies between 85 and 585 m asl. For some of the 15 settlements mentioned before 1150 AD, labels with place name and the year of the primary source is shown. (c) Landscape in the study area, view on Lindlar-Hartegasse.
<table>
<thead>
<tr>
<th>Accessibility measure</th>
<th>measures</th>
<th>Advantages (+) and disadvantages (–)</th>
</tr>
</thead>
</table>
| Degree of the node    | Attractiveness of a node in a graph | (+) can be calculated easily  
(–) focus on the local structure of the graph; weights of the edges and importance of the node are not taken into account; a high value does not mean that all directions are covered. |
| Degree of nodes in focal mobility networks | Attractiveness of a raster cell | (+) scale of accessibility can be adjusted by setting an appropriate cost limit for the ACS.  
(–) a high value does not mean that all directions are covered; complex calculation; drain procedure does not necessarily identify LCPs; no standard software available. |
| Shimbel index         | Attractiveness of a node in a graph | (+) can be calculated easily  
(–) focus on the global structure of the graph, edge effect; weights of the edges and importance of the node are not taken into account; does not allow the comparison of networks with different numbers of nodes. |
| Least-cost Shimbel index | Attractiveness of a node in a graph | (+) can be calculated easily.  
(–) focus on the global structure of the graph, edge effect; importance of the nodes is not taken into account; does not allow the comparison of networks with different numbers of nodes. |
| Shimbel grid index    | Attractiveness of a raster cell | (–) depends largely on the size of the study area; edge effect; high computational load. |
| Closeness centrality = integration | Attractiveness of a node in a graph | (+) can be calculated easily  
(–) focus on the global structure of the graph, edge effect; importance of the nodes is not taken into account. |
| Accessibility within a circle | Attractiveness of a raster cell / general accessibility | (+) scale of accessibility can be adjusted by the radius value; limited computational effort; limited edge effect  
(–) small areas of high cost values result in high accessibility values; undue importance attributed to locations close to the circle boundary. |
| Counting LCPs traversing a cell | Attractiveness of a raster cell / general accessibility | (+) scale of accessibility can be adjusted by setting a predefined distance from the initial point; limited edge effect  
(–) does not reflect the ease of reaching the cell; difficulties when more than 8 neighbours are considered in the raster grid; targets in small areas of high cost values result in high accessibility values. |
| Size of the site catchment | Attractiveness of a raster cell / general accessibility | (+) scale of accessibility can be adjusted by setting a predefined least-cost distance from the initial point; limited edge effect  
(–) undue importance attributed to locations close to the catchment boundary. |
| Accessibility to the nearest opportunity | Emissiveness of raster cells | (+) can be calculated quickly and easily  
(–) importance of the opportunities is not taken into account; only the nearest opportunity is considered. |
| Potential measure     | Emissiveness of a node in a graph or a raster cell | (+) takes all opportunities and their importance into account  
(–) with negative parameter w, result depends on the unit of measurement used for distance calculations and the result is undefined at the location of each opportunity; difficult to justify choice of parameters; edge effect. |
| Least-cost KDE         | Emissiveness or attractiveness of a raster cell / general accessibility | (+) takes all opportunities and their importance into account; scale of accessibility can be adjusted by setting the bandwidth.  
(–) no standard software available |

Table 1. Overview: measures of accessibility
Table shows for each accessibility measure the specific variant of accessibility calculated as well as the advantages and disadvantages of each measure.

3. Degree of Nodes in Focal Mobility Networks

Llobera et al. (2011) propose a method for extracting a graph from a cost grid with the aim of investigating the accessibility of a location. They first create an accumulated cost surface (ACS) based on this location. A hydrological procedure is applied to identify potential paths to the origin. The term focal mobility network is introduced for this set of paths. The authors mention the fact that alternatively the network could be derived by morphometric feature extraction. According to the definition of Llobera et al., accessibility of the location considered depends on the number of distinct paths within the network. This definition corresponds to a standard accessibility measure for graphs known as the order or degree of a node (Rodrigue et al. 2009, 29, 70). A diagram depicting the number of potential paths with increasing distance to the origin, serves to compare the accessibility of two locations. An extension of the method allows origins that cover several cells, i.e. lines or polygon shaped objects. Llobera et al. (2011) note that for polygon shaped origins, two or more focal mobility networks with different entry points may result. Extending this approach in order to assign an accessibility value to every cell in the raster map probably is computationally too expensive.

This method is only valid, if the resulting network really presents the most likely paths towards a given destination. It is well-known that hydrological drain procedures applied to the ACS do not necessarily identify the LCPs (Smith et al. 2007, 145–46). Moreover, counting the number of paths may not be appropriate if there are many paths to the south but no path in the north direction. It seems that the sizes of the obstacles separating the paths in the reconstructed network are not relevant for this accessibility definition. So with this approach, accessibility is based on morphometric features rather than on the actual friction values.

4. Shimbel Index and Shimbel Grid Index

In a graph, a simple accessibility measure is provided by the Shimbel index, which calculates the total number of edges needed to connect any one place with all the other locations (Waugh 2000, 615; Li and Shum 2001; Rodrigue et al. 2009, 70). The lower the Shimbel index, the higher is the accessibility of the location considered. This approach takes only the topological structure of the network into account, and this is inappropriate for raster data converted to a graph as explained above.

The least-cost Shimbel index calculates for each node the total path length (or path cost) instead of the edge count (Li and Shum 2001). This measure of accessibility can be applied to a cost grid converted to a graph (Rodrigue et al. 2009, 71). The term Shimbel grid is used for a raster accessibility map that stores the least-cost Shimbel index for each cell in a raster grid. A Shimbel grid for n raster cells requires the creation of n ACS grids. So this method is computationally expensive, and the effort increases very fast with increasing n.

The Shimbel grid index is sensitive to changes in the size of the study area. The path costs to places very far from the location considered are included in the calculation though they are probably not important at all for this location. Moreover, locations close to the border of the study area are often assigned a low accessibility value, as the costs of reaching them are high for most other locations in the study area. This edge effect is illustrated by the examples given by Rodrigue et al. (2009, 69).
5. Closeness Centrality or Integration

The least-cost Shimbel index is closely related to the index of closeness centrality used in network analysis for calculating the centrality of a node (Newman 2005). The index of closeness centrality is the average least-cost distance of a node to all other locations in the network, i.e. the least-cost Shimbel index is divided by the number of nodes in the network. This normalization enables the comparison of networks with different numbers of nodes. Extending on space syntax, this key figure has also been applied for analysing urban DEMs: Ratti (2005) uses the term integration for the “average length of the journeys to go to all other points”. Richards-Risetto (2012) calculates integration values on several scales for Maya sites in the Copán valley, Honduras, to derive results on the social interaction between and within the communities.

Low values of this index signify high centrality or accessibility, and therefore often the reciprocal of this index is used instead, so that high values indicate high centrality. Instead of the reciprocal, sometimes a simple transformation like \( f(x) = \frac{1}{x+1} \) is used to avoid division by zero. Such a transformation is problematic because the units of measurement influence the outcome of the index calculation. This cannot be fixed by multiplying the index by the appropriate factor for converting from one measurement system to another.

6. Accessibility within a Circle

Llobera (2000) defines the accessibility for a cost grid raster: For each cell the average path cost of moving from any cell within a certain radius to the selected cell is calculated on the basis of an ACS. The negative value of this average is the accessibility index. Llobera proposes adding an appropriate constant to ensure that positive accessibility values result. But adding a constant is also a problematic transformation with respect to changing the units of measurement. The approach by Llobera (without adding a constant) is in fact the index of closeness centrality restricted to the cells within a certain radius of the initial location and with another (better) way of dealing with the fact that a low sum of cost distances corresponds to high accessibility.

The radius \( r \) can be set to a small value if accessibility is to be analysed on a local scale (\( r = 100\text{m} \)), at a middle range (\( r = 1000\text{m} \)) or to a larger value for a global scale (\( r = 6000\text{m} \)). The radius values given in parantheses were suggested by Llobera. So this method allows accessibility calculations on different scales. Compared to the Shimbel grid, this approach suffers less from the problem that locations at the border of the study area receive a low accessibility value. Moreover, this approach is less computationally intensive. The main disadvantage of Llobera’s definition of accessibility is an inherent drawback of averaging (Fig. 2).

7. Counting LCPs Traversing a Cell

Whitley and Burns (2007) derive travel corridors from LCPs connecting a series of points spaced at regular intervals around the entire perimeter of the study area. According
Calculating Accessibility
Irmela Herzog

725
to the definition of Whitley and Burns, the accessibility of a raster cell is then given by the number of LCPs passing this cell. By applying this approach to a circular study area, the selected points on the perimeter will receive constant high accessibility values, whereas the LCP count for all other points on the perimeter probably will be zero.

Verhagen (2010) proposes an approach based on the ideas of Whitley and Burns, that also includes the scalability of Llobera’s method: A grid of points with a distance of 1 km is placed in the study region, and for each point 72 LCPs are constructed connecting the point with 72 targets radially distributed at a predefined distance from the initial point. This predefined distance determines the scale of the analysis. Verhagen uses distances of 250, 1000, 2500 and 5000 metres. As with the approach of Whitley and Burns, the accessibility of a cell is derived from the cumulative cost path map that counts the number of LCPs passing through each cell.

Counting the number of LCPs passing through each cell is not straight-forward if the number of neighbours considered during the raster to graph conversion exceeds 8. Moreover, with these approaches the costs of reaching two locations with identical LCP counts may vary considerably. For example, when the costs of movement in study area A are twice as high as in area B, both approaches assign identical accessibility values to the cells in both study areas, though intuitively higher values in area A are expected. Therefore, with both definitions of accessibility, the accessibility value of a location does not reflect the ease of reaching this location.

8. Size of the Site Catchment

Accessibility can also be measured in terms of the size of the site catchment (Conolly and Lake 2006, 214–15). The site catchment is a least-cost buffer, and in the study region, it is the area included by the isoline corresponding to a certain radius of movement on level and dry ground (Fig. 3).

Figure 3 shows the site catchments on different scales for the two settlements Wipperfürth and Wiehl. Both Wipperfürth and Wiehl were first mentioned in 1131 (Pampus 1998), and both places are situated along rivers in east-west direction. But Wipperfürth

![Figure 3. Site catchment bands for two settlements in the study region: (a) Wipperfürth; (b) Wiehl. Choosing the maximum cost limit for generating the catchment is not easy. The inner catchments are generally more important than the bands at a distance. For both settlements, the width of each band corresponds to the costs of moving 1 km on level and dry ground. A comparison of the catchment sizes can be found in Table 2.](image)
played a more important role in Medieval times because it was at a crossroads of several ancient trade routes and the first settlement in the study area that received town privileges (in 1217). The comparison of the site catchments for these two settlements shows that the accessibility of Wipperfürth is higher than that of Wiehl (Tab. 2).

The site catchment accessibility concept is based on the assumption that all locations within the site catchment are of equal importance. But in the study area, the catchment of a Medieval settlement normally consisted of several successive zones, i.e. bands with gardens, orchards, meadows, and fields (Nicke 1995, 52) that needed different intensity of attention by the inhabitants of the settlement.

Site catchments delimited by isochrones are equivalent to potential path areas as defined in time geography (Mlekuž 2013): A basic concept of time geography is the space-time path that traces the movement of an individual in space and time (Kim 2003), and the set of reachable locations forms the potential path area. According to Kim, the size of the potential path area has been used as a measure of accessibility in time geography.

Mlekuž (2013) suggests adding site catchments for all cells in the raster map to generate a cumulative potential path area resulting in a general accessibility map. Mlekuž notes that generating this map is computationally expensive and suffers from the boundary effect.

<table>
<thead>
<tr>
<th>Catchment radius</th>
<th>Wipperfürth area (sq km)</th>
<th>Wiehl area (sq km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.54</td>
<td>0.32</td>
</tr>
<tr>
<td>2</td>
<td>3.16</td>
<td>1.90</td>
</tr>
<tr>
<td>3</td>
<td>8.19</td>
<td>4.80</td>
</tr>
<tr>
<td>4</td>
<td>15.78</td>
<td>10.34</td>
</tr>
<tr>
<td>5</td>
<td>26.09</td>
<td>18.96</td>
</tr>
<tr>
<td>6</td>
<td>38.72</td>
<td>29.30</td>
</tr>
<tr>
<td>7</td>
<td>54.92</td>
<td>42.90</td>
</tr>
<tr>
<td>8</td>
<td>73.63</td>
<td>47.91</td>
</tr>
</tbody>
</table>

Table 2. Comparison of the site catchment sizes of Wipperfürth and Wiehl (Fig. 3). A catchment radius of 1 corresponds to the costs of moving 1 km on level and dry ground. On every scale considered, the catchment size of Wipperfürth exceeds that of Wiehl.

Figure 4. Maps showing the accessibility to the nearest opportunity for sources (left) and for ancient trade routes (right) in the north of the study region (black frame in Fig. 1). Low accessibility values are shown in dark grey, white signifies optimal accessibility.
9. Accessibility to the Nearest Opportunity

The European Union (1995-2010) created an accessibility map that shows the travel costs from each cell in the study area to the nearest target location. This method can be implemented with an anisotropic cost model if both the route to the opportunity and back are taken into account. The computational load is reduced significantly by computing the minimum travel costs starting from all opportunities, i.e. reversing the process. In general the number of opportunities is quite small compared to the total number of raster cells. This reversed algorithm is probably the fastest way for creating a reasonable accessibility map for a given set of opportunities of equal weight. Note that high values in such an accessibility map correspond to low accessibility and vice versa. Transformations resulting in the opposite relationship have been discussed above.

In a GIS, such an accessibility map can be calculated easily by creating simultaneous least-cost buffers for all opportunities. A spreading process is started at each opportunity cell, and stops whenever two buffers meet at a point with equal least-cost distance to two opportunities (Fig. 4). The approach can be implemented for all opportunity shapes. The complexity of this calculation corresponds to that of calculating an ACS.

Such accessibility maps can be used for statistical tests. For example, Nicke (1995, 49) presents two hypotheses that could be tested this way: (i) The earliest settlements in the study region preferred locations close to the ancient roads; (ii) in a later phase new settlements developed close to the sources. Testing such hypotheses is one of the oldest GIS applications in archaeology (Kvamme 1992). However, Kvamme used straight-line distances to the opportunities whereas the accessibility maps presented in Fig. 4 rely on calibrated least-cost distances. The background distance-to-opportunity-distribution is derived from the accessibility values of all cells in the map. This distribution is compared to the distribution of the accessibility values of the known Medieval settlements by applying a Kolmogorov-Smirnov goodness-of-fit test. According to the test, the 513 settlements in the study area first mentioned before 1500 are not significantly closer to the ancient roads than random points (Herzog in press).

According to Miller (1999), accessibility to the nearest opportunity is an example of a maxitive accessibility approach, i.e.

\[ f(d_1, a_1), ..., (d_n, a_n) = \max\{z(d_1, a_1), ..., z(d_n, a_n)\} \]

where \( n \) is the number of opportunities, \( a_k \) is the weight of opportunity no. \( k \), \( d_k \) is the (cost-) distance between the location considered and opportunity no. \( k \), where the function \( z \) represents the utility of the opportunity. For the accessibility to the nearest opportunity approach, the utility of an opportunity depends only on its distance, its importance is not included in the calculation.

Moreover, only the nearest opportunity is taken into account. But the presence of several opportunities in the neighbourhood is helpful, if one of the opportunities does not serve its purpose for some time, e.g. the source of water runs dry or is polluted. It is well-known in modern economies that customers often benefit from several competing opportunities, whereas monopolies should be avoided. Additive accessibility approaches as described by Miller (1999) are probably more appropriate for most archaeological purposes.

10. Potential Measure

The general formula for calculating additive emissiveness is

\[ f[(d_1, a_1), ..., (d_n, a_n)] = z(d_1, a_1) + ... + z(d_n, a_n) \]

where \( n, a_k, d_k \) are defined as above.
Fotheringham et al. (2000, 245) present a commonly used additive emissiveness formula called potential measure:

\[ \text{Accessibility} = \sum_j a_j d_j^w \]

So the utility function \( z \) is: \( z(d,a) = a^v d^w \), \( v \) and \( w \) reflect the importance of the weight and distance respectively. The default values for \( v \) and \( w \) are: \( v = 1 \) and \( w = -1 \).

This approach takes opportunity weights into account. Opportunity weights might be measured in terms of square metres, quality and quantity of some desired substance available at the given location etc. Li and Shum (2001) point out the difficulties in identifying a variable that could serve as weight in modern networks where many variables are known. For networks in prehistoric times estimating weights is even more difficult.

With negative \( w \), nearby opportunities contribute more to the accessibility than those at a distance. This effect can be emphasized by increasing the absolute value of \( w \): In a gravity model (Waugh 2000, 618), \( w \) is set to \(-2\). By decreasing \( w \) (e.g. \( w = -4 \)) the accessibility surface becomes more ‘spiky’. Fotheringham et al. note that there usually is no justification for the choice of constants \( v \) and \( w \), except that the results appear reasonable.

If the focus of the analysis is on the presence or absence of certain opportunities, then all \( a_j \) should be set to 1. When every opportunity has weight 1 and \( w \) is set to 1, this accessibility measure is identical to the least-cost Shimbel index presented above.

The accessibility value increases when enlarging the study area so that additional opportunities are included. The following example illustrates a drawback of this concept: A typical opportunity in a prehistoric context is a fresh water supply. The weight in this case is the amount of water supplied. If the source is 100 km away but produces 1000 times the daily need of the community considered, the contribution of this source to the accessibility measure with default parameters is 10 times as high as the source that is 1 km from the starting location and provides just adequate water supply. Even if \( w \) is set to \(-2\), and \( v = 0.5 \), so as to reduce the influence of high-weight and distant opportunities, the source at the far distance still contributes 3.16 times as much to the accessibility value as the nearby source. This example shows that a maximum distance should be set as in the method considering accessibility within a circle.

For all negative values of \( w \), the formula is undefined if \( d_j = 0 \) for some \( j \) (Fig. 6). When an accessibility map is to be created, this means that the map is undefined at all opportunities. For this reason, this popular accessibility measure violates the third axiom of Weibull (Miller 1999), that requires that the accessibility measure is continuous and increasing for \( d_j = 0 \) and increasing a \( j \). Rodrigue et al. (2009, 71–72) try to fix this problem by setting the accessibility value to \( a_j^v \) for \( d_j = 0 \). However, if \( d_j < 1 \), then the accessibility value of this nearby point exceeds that of the opportunity location. When calculating accessibility in a raster grid, the distance units may be chosen so that the distance of any two neighbouring raster cells is larger than 1. But still the resulting accessibility map is not continuous.

The function \( z(d,a) \) may also be considered as a decay function, which models the decreasing influence of the opportunity with increasing distance. A decay function depending on the distance from an opportunity was also proposed by Ratti (2005) resulting in a weighted catchment of the opportunity. A semi-bell-shaped Gaussian decay function seems to be more realistic than the function of the accessibility formula presented by Fotheringham et al. (Fig. 6). For example, the bell shape is reflected in an ethnographic study from Nigeria investigating the number of trips per person depending on the distance to the neighbour visited (Davis Stone 1993).
Calculating Accessibility
Irmela Herzog


The Gaussian bell-shaped function is often chosen as the kernel in the formula of kernel density estimation (KDE), and the formula for KDE bears some similarity with the general additive accessibility approach if all weights are identical (Scott 1992, 125–55). KDE can be computed on the basis of different decay functions, and a bounded kernel (i.e. that is zero beyond a predefined maximum distance) is less wasteful in terms of computation time and limits the edge effect. My least-cost KDE program (Herzog and Yépez 2013) implements the radial-symmetric Epanechnikov kernel because it can be calculated easily and is bounded (Fig. 6).

The size of the kernel is determined by the spread parameter or bandwidth, that is dependent on the predefined maximum distance for bounded kernels. The user should select the bandwidth according to his or her knowledge of the movements of the prehistoric people considered, i.e. the maximum cost in energy or time expended for daily movement or for travelling at a larger scale. Weighted kernels can be used to account for opportunity weights (Smith et al. 2007, 131). When choosing a bounded kernel, important opportunities do not dominate large regions as with the potential measure, but the radius of influence remains restricted. In some cases, it might be desirable to have a larger radius of influence for important opportunities, and this is modelled in a straight-forward way by choosing different bandwidth parameters. The concept of KDE can be extended to least-cost KDE by applying a radial-symmetric basic kernel and replacing the straight-line distance by the least-cost distance (Herzog and Yépez 2013; Fig. 5). The least-cost KDE accessibility measure is an additive accessibility measure where

\[ z(d,a) = a*K(d,b), \]

and K is a one-dimensional kernel function (applied radial-symmetrically); the bandwidth parameter b controls the scale of the analysis.

Least-cost KDE can also be regarded as a refinement of the accessibility concept based on site catchments. The refinement takes into account that the area close to the central location is more important than an area near the catchment boundary. The catchment area is subdivided into adjacent bands delimited by equidistant isolines; the bands are assigned...
decreasing weights starting at the central location. The corresponding kernel exhibits a stepwise decay and its shape can be compared to a step pyramid with an irregular base. Decreasing the size of the steps results in a smoother kernel.

The least-cost kernel approach can also be adjusted to line-shaped opportunities. Note that converting a line-shaped opportunity to raster cells covering this line and calculating the accessibility map for these raster cells (with the help of least-cost kernels) is not appropriate because this results in high accessibility values for central points of the line but not at the ends. The correct method requires calculating the ACS (delimited by the bandwidth) for each line. Afterwards, for each ACS the values are modified according to the Epanechnikov-kernel function:

\[ f(x) = a (1 - (x/b)^2) \text{ for } x < b, 0 \text{ otherwise} \]

where \( a \) is the importance of the line-shaped opportunity, and \( b \) the maximum distance. The accessibility map is the sum of these transformed ACS grids. Examples of least-cost KDE accessibility maps for point- and line-shaped opportunities are shown in Fig. 7.

Kolmogorov-Smirnov goodness-of-fit tests can be applied for such maps as well. If the null hypothesis is rejected, there is a close relationship between the location of the sites and the opportunities. In contrast to the nearest opportunity approach, the KDE accessibility maps take all neighbouring opportunities and their importance into account.

Moreover, the KDE approach can be used to generate general accessibility maps.
by applying the method on equally weighted point opportunities in a grid (Fig. 8). The grid resolution should be well below the bandwidth parameter (10% or less). With high grid resolutions the computational load is quite heavy. The example in Fig. 8 shows that the ancient trade routes in the study area are mainly located in areas of high general accessibility. A Kolmogorov-Smirnov goodness-of-fit test could be applied in this case as well to confirm this hypothesis. Due to the edge effect the test should be restricted to the area within the inner frame.

Least-cost kernels can also be interpreted within the conceptual framework of time geography if a cost function measuring time is chosen. In time geography space is represented by two coordinates and time is shown on the z-axis (Kim 2003). A space-time prism in the simplest case is a cone that is centred on the initial location of the person considered and shows which locations can be reached within a certain time limit, and the volume of the space-time prism is considered as a measure of accessibility (see also Miller 1999). This space-time prism is equivalent to the least-cost kernel with a basic kernel shape of a cone. If the attractiveness of an opportunity is to be considered, this cone can also be viewed as a visualisation of the difference between the utility of that location and the required cost for travelling from any point in the study area to the opportunity (for positive differences).

An alternative kernel is the generic utility function presented by Miller (1999), but it is difficult to estimate the parameters of this function. In any case, this kernel decreases rapidly close to the location of the opportunity, and this is a substantial drawback. Kim (2003) notes that nobody will travel for 2.5 hours in order to participate in a discretionary activity for just one minute. So the locations that can be reached well before the time limit is expended are more important than those on the outer isochrone. For this reason, the Epanechnikov kernel is more appropriate than the cone. By applying the least-cost Epanechnikov kernel a probabilistic model is created, that assigns higher probabilities to movements in the vicinity of the origin of movement.

In sum, least-cost KDE can be used to calculate both general accessibility and accessibility with respect to opportunities in a space-time framework. However, there is an ongoing debate whether time or energetic expenditure is the limiting factor determining patterns of movement in past times, i.e. if it is appropriate to apply the concepts of time geography in a prehistoric or Medieval context (e.g. Kantner 2012).

This section has shown that least-cost KDE provides a versatile and useful tool kit for accessibility calculations: It allows to model opportunities of different importance as well as point-shaped and line-shaped opportunities. Moreover, when applied to an appropriate set of raster points, the end-result is a general accessibility map. The choice of the bandwidth parameter provides an intuitive way to control the scale of the calculations. With a bounded kernel like the Epanechnikov kernel, the results are local, i.e. opportunities at a large distance are not included in the accessibility value.

12. Future Work

For simplicity and due to its optimality in statistical terms, the Epanechnikov kernel was implemented and applied for all least-cost KDE results shown. However, the first derivative of this kernel is not continuous at the boundary of the support interval, and it does not come as close to the bell shape as the cos(x)+1 kernel, that has been applied successfully to estimate the density of simulated point patterns (Herzog 2007). It is planned to add the capability of choosing different kernels, this will allow to adapt the least-cost KDE method to ethnographic evidence.

Another future task is to extract natural pathways from the general accessibility map.
created by KDE methods. The cells with the highest accessibility values often form disconnected patches instead of proper route networks. Moreover, this simple approach will not identify probable paths in large areas of low accessibility. To address this issue, Mlekuž (2011) proposed applying Llobera’s prominence principle. Llobera (2001) defines prominence for each point of a digital elevation model in relation to the local surroundings so that he is able to identify peaks and ridges. The scale of the calculation can be controlled by the radius of the surroundings considered. Applying this approach to general accessibility maps allows to locate probable paths that correspond to the ridges in the accessibility map.

Acknowledgements

I would like to thank the two anonymous reviewers for their useful hints and recommendations.

References


Wünsdorf: Druckhaus Köthen.


Simulated Paths, Real Paths? A Case Study of Iberian Cessetania (Iron Age Society)

Joan Canela Gràcia
Institut Català d’Arqueologia Clàssica (ICAC), Spain

Abstract:
Cessetania was an early Iberian state in the north-east of the Iberian Peninsula. Early states appeared in this area as a result of the last Bronze Age societies evolving into the more complex hierarchical settlement types that developed from the 6th to the 2nd century BC. Our aim is to study possible ancient paths in Cessetania. In order to do this, we have chosen three important Iberian settlements and have attempted to calculate LCRs (Least Cost Routes) using different methods. We have sought out coincidences with real historical paths (most of which we know from the Middle Ages). Our purpose is, on the one hand, to prove the plausibility of the simulated paths and on the other to investigate the possibility that these historical paths could be reminiscent of ancient paths.

Keywords:
Least Cost Routes, Cessetania, Protohistory, Ancient Paths

1. Introduction

This paper presents a preliminary study into how research into communication routes in the protohistoric period can be focused. It is not, therefore, a definitive study. In studying the ancient paths and communication routes in Cessetania (Fig. 1) we come up against some important impediments. The landscape has undergone major changes in the last 2,500 years and the protohistoric traces appear to have been obliterated or greatly modified by later phases, with the most drastic modifications having come about in the last fifty years. We take as our thesis that for commercial, social and political reasons the most important settlements would have developed communication routes between themselves and that these would have become links between the coast and the interior. We have chosen three settlements in our area of study that can be considered as “urban”. The first is Cesse; the site is now within the urban fabric of the modern city of Tarragona and is considered by several experts to have been the capital of the archaic state of Cessetania (Asensio et al. 2001, 253-271). Located in the lower part of the present-day city, it covered an area of approximately seven hectares. The second is El Vilar (Valls, Tarragona). Located on the route between the coast and the interior, it was the capital of its surrounding area and would have occupied approximately six hectares (Fabra and Vilalta 2008, 163-185). The third is Les Masies de Sant Miquel (Banyeres del Penedès, Tarragona), which covered an area of at least three hectares (Cela et al. 2003, 25-264) and played a similar role to that of El Vilar, but in the El Penedès region, a natural inland corridor to the north that avoids the rocky sector of the Garraf coast (Fig. 2).

The territory between these three centres coincides with the central zone of Cessetania, which has a typical Mediterranean relief where, in a relatively small geographical area, there are large sectors of plains and mountains that extend to the sea. The two main water courses are the Francolí and Gaïà Rivers with low volumes of water that flow into the sea in Tarragona and Tamarit respectively. The main mountain sector is the Serra del Montmell with peaks almost a thousand metres high. They act as an orographic barrier between the interior of the Camp de Tarragona and El Penedès, although there are many passes that allow them to be crossed (Fig. 3).

Taking these settlements as starting points, we calculated the respective LCRs (Least Cost Routes) based on four different
algorithms. Subsequently we compared them with ancient documentation and cartography to look for possible coincidences with historical roads. The final objective was to compare the plausibility of this type of calculation using GIS (Geographic Information Systems) and also to investigate the possibility that the historical roads may be older.

2. Methodology

We used ArcGIS 10 software for these calculations. The first step was to choose the DEM (Digital Elevation Model) for calculating the LCRs. Trials were carried out with the Catalan Cartographic Institute’s 15x15 DEM and the Spanish National Geographic Institute’s 5x5 DEM, but noise clearly affected the results, with modern infrastructures often diverting the LCRs. The DEMs had to be processed and refined so that we could work with them. We decided to carry out a trial with a newly-created 5x5 DEM based on information from the Catalan Cartographic Institute’s 1:5000 topographic base edited as a point cloud layer. This allowed us to easily modify the noise, eliminating those that formed part of modern infrastructures or altering their characteristics. We edited the point cloud to remove roads, motorways and railway lines. In this way we obtained the basis for reproducing a DEM more in keeping with the ancient landscape and without interference from modern structures that strongly influence the calculated LCR trajectories. This is the easiest way of generating the most suitable DEM for our calculations and is preferable to modifying predefined DEMs such as those mentioned above. Once the point cloud had been converted to a DEM using the Topo to Raster tool, we created the working 5x5 DEM (Fig. 4).

Research on LCR has been in development for many years (Bell and Lock 2000; Bell et al. 2002; Bellavia 2006; Fiz 2008; Fiz and Orengo 2008; Gietl et al. 2008; Herzog 2010; Herzog and Posluschny 2008; Howey 2007; Llobera and Sluckin 2007; Murrieta 2010; Zakšek et al. 2008). There are many formulas for calculating LCRs and to begin we decided to use four classic algorithms that focus on the question of movement from different perspectives, both isotropic and anisotropic. Hiker’s and the
Effective Friction algorithms are only based on the slope, whereas Pandolf and Van Leusen’s take other variables, such as the energetic cost of the transport and the individual, into account. They are the following:

- **Hiker** (Gorenfo and Gale 1990, 240-274)
  
  \[ V = 6e^{(3.5s + 0.05)} \]
  
  Where:
  
  \[ V = \text{Velocity} \]
  \[ s = \text{slope} \]

- **Silva & Pizziolo** (Silva and Pizziolo 2001, 279-286)

  \[ \text{Effective Friction} = \text{stated friction} [f] \]

  Where:
  
  \[ f = \cos k \Delta \alpha \]
  \[ k = \text{user defined coefficient (foot movement)} \]
  \[ k = 2 \]

  \[ \Delta \alpha = \text{difference angle between the direction of movement that incurs the maximum friction and the direction of movement being considered.} \]

- **Pandolf** (Pandolf et al. 1977)

  \[ M = 1.5W + 2.0(W+L)(L/W)^2 + N(W+L)(1.5V^2 + 0.35V^* \text{abs}(G+6)), \]

  Where:
  
  \[ L = \text{Carried weight (4 kg)} \]
  \[ V = 5\text{km/h} \]
  \[ N = \text{Terrain factor (rivers)} \]
  \[ W = 60\text{ kg} \]
  \[ G = \text{Slope} \]

- **Van Leusen** (Van Leusen 2000)

  \[ M = 105.483 + 255.44N + 124.32NG \]

  Adapted algorithm from Pandolf where:
  
  \[ W = 70\text{ kg} \]
  \[ L = \text{Carried weight (4kg)} \]
  \[ V = 4.8\text{km/h} \]
  \[ N = \text{Terrain factor (rivers)} \]
  \[ G = \text{Slope} \]

Research into movement was carried out using two different types of algorithm - isotropic and anisotropic. The former is based on the fact that any movement involves a cost without the direction being definitive (Wheatley and Gillings 2002, 151), whereas in the latter the direction is essential to facilitate or complicate the movement (De Silva and Pizziolo 2001, 279). Common variables in the former are the luxuriance of the vegetation or the type of ground, assigning a significant quantity to the transit for those topographical units. In the latter, the factors that determine, facilitate and hinder movement include gradient and hydrographic flows. These algorithms are used in GIS to calculate the cost of translocation for each of the settlements studied, an intermediate step towards finally calculating which are the paths with the optimum cost compared to the rest (Fiz 2008, 203-204). In this initial experiment we chose to respect the original formulas and not to assimilate parameters in order to see whether they gave rise to very different results, although obviously in the future these parameters will have to be homogenised.
In order to compare the results with the historical paths, we made an initial data collection using the first National Geographic Institute 1:50000 topographic series map and the so-called “American Flight of 56/57” (aerial photographs taken by the United States Air Force in 1956-57) (Fig. 5). These are easily accessible from GIS as WMS (Web Map Service). We are also compiling documental sources (especially from notarial archives), ancient cartography (from the Peninsular War onwards) and books about paths and roads. The documental sources are particularly useful as in some cases they mention long-distance communication routes, especially in capbreus (a kind of land inventory) or documents of a similar type. We also used maps from the beginning of the twentieth century, the 1:10.000 municipal maps made by the Mancomunitat de Catalunya. With all these sources we created an historical path layer at ArcGIS.

Once we had obtained the results of the LCRs, we compared them to the information we had compiled previously on the territory’s historical paths. We noted any coincidences found.

### 3. Results

As mentioned, this is a work in progress and in this article we present the basic practical and methodological proposals of our study. We will continue to work on adjusting the LCR calculations to include factors such as hydrology, which we have not taken into account in this initial trial, but that we consider necessary.

The results (Fig. 6) were uneven in each model; in some cases we found many matches and in others hardly any. We also believe that the DEM we used was not of sufficient quality and that the same calculations made with another with higher resolution DEMs (for example the National Geographic Institute’s 2X2 model) would provide improved results. The most obvious case is the Pandolf algorithm, which ended up generating a result that was too rectilinear.

Part of the ancient coastal route of the Via Augusta (the modern N-340 highway) was followed approximately by all the models, with the Effective Friction algorithm following it the
longest (B). The routes calculated between El Vilar and Cesse contributed many more matches; only Hiker and Van Leusen corresponded in part with the path from Valls to Vallmoll (A). We found more points in common for the route between El Vilar and Les Masies de Sant Miquel. Particularly interesting was the match in the place where the River Gaià was crossed by the path from Montblanc to Vilafranca (D), which some researchers have identified as the interior route of the Via Augusta towards Ilerda. Also worthy of mention is the match of Hiker, Effective Friction and Van Leusen in the point of the first crossing of the torrent, coinciding with the Valls to Alió route (E). The models that gave the best results were Hiker, Effective Friction and Van Leusen, while Pandolf gave discrete results. In our opinion it is too early to come to any conclusions about which models work best in which situations, as we are still at an incipient stage of the research; however we believe we are opening up an interesting line of research in this field.

4. Conclusions

As tends to happen with such investigative processes, the new research, instead of offering solutions, often poses more questions for which answers have to be found.

It is hard to find the right point when it comes to calculating the LCRs. It is easy to end up forcing the formula to the extent that the calculated route coincides with our working hypotheses. For example, if we wish to calculate a route and make it fit our hypothesis, we can adjust the different calculation factors to obtain the desired result by forcing the formula.

It can also be observed that on high road routes, elements (e.g. a stream) that in some places hinder passage, in others act as corridors, facilitating communication (e.g. torrents). We wish to take into account all those factors that influence movement (slope, direction, hydrology, etc.), but without “subjectivising”
the model in such a way that it ends up giving us a result we like, but that is not necessarily correct.

Another important question is whether the coincidence between an LCR and a historic path is really showing us an ancient (protohistoric) route. The study of communication routes is extremely difficult and complex and due to their erosive nature it is not always possible to date them archaeologically. Dating is more reliable from the Roman, mediaeval and modern periods. LCRs and their comparison with historically documented paths could be a good approximation methodology, but fieldwork is needed to corroborate these hypotheses, with surveys and archaeological excavations at places along these routes that could provide us with information about their use over time.

References


Simulated Paths, Real Paths? A Case Study of Iberian Cessetania (Iron Age Society)
Joan Canela Gracia


Open Source GIS for Archaeological Data: Two Case Studies from British and Egyptian Archaeology

Anna Kathrin Hodgkinson
University of Liverpool and Oxford Archaeology North, UK

Luca Bianconi and Stefano Costa
IOSA Project, Italy

Abstract:
Proprietary software can put archaeological data at risk by placing unnecessary barriers of entry to studying the past through closed file formats. Open source software has been successfully adopted for visualisation and analysis of archaeological data in both commercial and academic archaeology. Two case studies are presented here: a large commercial fieldwork project in England and an archaeological and topographic survey at the site of Medinet el-Gurob in Egypt. Open source software - as we describe - replaces proprietary software both within the highly cost-sensitive commercial sector and within international and educational archaeology. This is done through active use and testing, as well as dissemination of documentation. Software discussed includes tools for digital survey download (Total Open Station), GIS (gvSIG, Quantum GIS (QGIS), GRASS, PostGIS), Illustration (Inkscape) and 3D analysis (Paraview, VisIt). The move from proprietary software to open source has been a success both within this sector of British archaeology and this Egyptian project.

Keywords:
Open source, GIS, Total Station, Egypt, Oxford Archaeology

1. Problems and Objectives

Proprietary software can place archaeological data under the threat of becoming "locked in", as it is usually saved in proprietary formats over which users have no control. For instance, the use of CAD software can present a threat for the archaeological data with its inability to maintain topology in vector data. The move away from CAD and other proprietary spatial software has formerly been criticised and declared problematic by fellow archaeologists, the main issue faced being the lack of documentation and training time. Archaeological site survey and GIS is still frequently considered a specialist task, inaccessible to the archaeologist fieldworker. The aim was to eliminate this thinking.

As commercial archaeology usually faces a large amount of time-pressure, it has been argued that teaching staff the use of open source GIS would be time-consuming. However, it became increasingly clear that having good documentation would lead to more cost effective training methods for staff as well as safer, more sustainable data archives overall.

2. Commercial Archaeology and Open Source Software - the Oxford Archaeology Case Study

In 2009, as part of a large-scale road scheme, a data visualisation project was set up for the fieldwork phase only using open source GIS software. The in-house package, gvSIG OA Digital Edition (OA DE) was selected as the desktop GIS package for data visualisation and initial analysis. For extended functionality QGIS was used alongside this. The gvSIG OA DE, based on gvSIG 1.10, has been adapted by the company in a way that would make it
Open Source GIS for Archaeological Data: Two Case Studies from British and Egyptian Archaeology
Anna Kathrin Hodgkinson, Luca Bianconi and Stefano Costa

It is easier to use, not only within archaeology, but also for external users: some program menus were restructured, installation was made easier and a number of GRASS GIS modules were integrated by Benjamin Ducke. In addition, the English language user interface was improved (gvSIG is originally developed with Spanish as its primary language). The GRASS GIS inclusion occurred at a later stage (in 2010) and integrated a number of functions into the SEXTANTE toolbox, which forms part of gvSIG. This greatly enhanced the software’s capabilities through almost eliminating the need for the user to understand and apply the often complicated GRASS setup and workflow, while enabling them to use the highly powerful GRASS commands often unavailable in desktop GIS applications. GRASS is often used to convert between different file and vector type formats, copes with CAD DXF data very well, and is powerful when it comes to 3D modelling and maintaining a clean topology. Being able to make customisations such as the ones behind the gvSIG OA DE is a freedom only granted by open source GIS software.

After successfully setting up a project on site, the data from previous phases of the same archaeological project was exported from CAD in ESRI shapefile format and thus linked to the GIS. The gvSIG OA DE was then used on a daily basis to import the downloaded survey data, to print maps, make database queries and georeference images, both hand-drawn site plans and hi-view photographs, which were produced on site on a daily basis. Some initial digitising was also undertaken (this usually being part of the post-excavation process).

During the excavation phase, databases were set up in Microsoft Access 97, as Oxford Archaeology’s migration to PostgreSQL had not yet taken place at the time. However, all archaeological databases, both those produced during the fieldwork and the post-excavation stages were later successfully migrated to a PostgreSQL database, complete with a spatial PostGIS component, which was visualised using QGIS.

Our on-site work-flow was documented in detail and released in the shape of an online

<table>
<thead>
<tr>
<th>Task</th>
<th>Old Workflow</th>
<th>New Workflow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site survey using Total Station or GPS and download</td>
<td>Geomatics specialists only</td>
<td>Any Archaeologist</td>
</tr>
<tr>
<td>Visualisation of archaeological survey data</td>
<td>AutodeskMap CAD and ESRI ArcGIS</td>
<td>Open Source GIS: gvSIG OA DE, Quantum GIS, PostGIS, GRASS Any Archaeologist</td>
</tr>
<tr>
<td>Illustration, figure production for client reports and monographs</td>
<td>Adobe Illustrator, InDesign etc. Graphics specialists only</td>
<td>Open Source Illustration software: Inkscape Any Archaeologist</td>
</tr>
</tbody>
</table>

Table 1. Survey and GIS manual, Hodgkinson 2010: On-site Survey.

Figure 1. Survey and GIS manual, Hodgkinson 2010: On-site Survey.
3. Gurob

The site of Gurob, also known as Medinet el-Gurob (Petrie 1890, Shaw 2011) is an ancient Egyptian settlement best known for its “Harem” palace, Temple, and vast cemeteries, although a settlement is known to have been in place as well. Its main period of occupation was during the Egyptian New Kingdom (c.1550 – 1070 BC, Late Bronze Age), although there is evidence for earlier settlement activity, such as cemeteries, dating as early as to the Predynastic Period. The temple and palace were founded by king Thutmose III (1479-1425 BC) and the site was used until the Ramessid period. The site is located at the entrance of the Egyptian Faiyum region, just outside the cultivated land, in the desert. The area occupied by the site measures just over 1.2km². It was first excavated by Flinders Petrie (1888-90), and later by Curelly and Loat (1904), Borchardt (1905), and Brunton and Engelbach (1920). No full plan was produced of the site, and little remained known throughout the best part of the 20th century, as the site used by the Egyptian military for training purposes from ~1959 until the 1980s. Unfortunately, this also meant that the archaeological remains suffered badly. While Petrie (1890, 33) still claims to have seen traces of settlement at Gurob, very little except the outline of the palace enclosure can now be seen by the naked eye. In fact, the first composite plan of the palace was produced by Kemp (1978), but only based on published material from previous work undertaken at Gurob.

The University of Liverpool, under the direction of Ian Shaw, began working at Gurob in 2005, when a survey of the site was set up. The aim of this mission is to produce a detailed manual (Hodgkinson 2010, Fig. 1), which in turn was tested on site and refined by using it as training material for staff. This manual both guides the archaeologist through the steps of setting up their survey equipment and also explains to them the downloading of the survey data including an introduction to gvSIG, and the digitising and production of maps (Fig. 2).

Until recently the illustration of of high-quality printable maps for official client reports took place exclusively in proprietary software, such as Adobe Illustrator, because appropriate cartography tools were not available in open source GIS software. However, a further methodology was developed, involving new, improved cartography tools in gvSIG and QGIS, which were combined with the powerful open source illustration software, Inkscape, and produced highly satisfying results, which have been used in official client reports. Therefore, the survey and GIS work-flow has been extended, and a second manual has been published focusing on high-quality map illustration and additional cartographic techniques available in QGIS and Inkscape (Robinson, Campbell and Hodgkinson 2011).
plan of the site using a Leica TCR 705 total station (Fig. 3), as well as the establishment of a pottery and small finds corpus through surface survey, in order to understand Gurob as a New Kingdom Egyptian settlement and Harem city. In 2010, the first small-scale excavations took place with Anna Hodgkinson (who, between 2009 and 2011 was also the key surveyor at Gurob) working within the “Industrial Area” to the north-east of the palace. A collaboration with the University of Copenhagen has existed since 2011, and larger-scale excavations have begun to take place at Gurob from 2012 onwards (Shaw 2012).

As not much is visible on the surface a variety of survey methods have been used over the years in order to detect any remaining archaeological structures. Geophysical survey was undertaken; in 2006, a Proton Magnetometry survey was done, however, due to the large amount of metal objects left behind by the military the results were distorted and anomalies were often obscured. This survey was undertaken in the palace and industrial areas. The latter showed two circular anomalies where on the surface many traces of burnt and vitrified mud-bricks are visible. These results have been integrated into the site GIS in order to combine them with the results from total station survey and hand-drawn excavation plans.

In 2011, after the Spring season was postponed due to the January Revolution, many instances of looting were recorded. These were mainly encountered on shaft tombs, but many test pits within the core site and against the palace enclosure were also discovered. The survey team recorded these instances using Leica Zeno hand-mapper devices (Fig. 4).

Artefacts such as clay figurines, pieces of faience and glass, jewellery and stone vessels are frequently occurring surface finds, which are surveyed three-dimensionally with the total station (Fig. 7). Any other visible features, both ancient and modern are surveyed as well, and the results are compared against what can be seen on satellite imagery and magnetometry results.

Only open source geospatial software is used for the downloading and maintenance of Gurob survey data: For the download of digital
survey data acquired through total station survey Total Open Station (TOPS) is used, and QGIS for data maintenance, visualisation and analysis. A topographic survey undertaken by Hannah Pethen in 2011-12, using both the Leica TCR 705 and a Leica Viva TS15 robotic total stations resulted in an interpolated digital elevation model of the desert surface, which showed remains of structures, and which was subjected to viewshed analysis in both QGIS and GRASS GIS. The result was visualised in other open source 3D packages, such as Paraview and VisIt.

Site plans were digitised in QGIS and exported as scaled vector graphics (SVG) and, together with illustrated sections, retouched in open source vector editing software Inkscape. It was concluded that great results could be achieved without the use of proprietary software, such as ESRI ArcMap and Adobe Illustrator.

4. Total Open Station

Compatibility between devices and formats has become a serious problem, especially with the spread of the wide usage of many types of electronic devices for collecting and communicating data during archaeological work. One of those devices is the total station, commonly used within archaeological fieldwork for a range of topographic tasks. It is not always easy to achieve a communication between various models of total station and different operating systems. For example, Linux or OSX users are deeply disadvantaged when it comes to accomplishing the most common tasks with a total station, namely downloading the recorded survey data from the device to a computer.

The general problem can be divided into two different aspects: one related to the data format and one of good practice. The first aspect is related to the different formats the data are saved as: there is no industry standard for this data and each vendor has its own implementation according to their own specifications, and thus different models of total station do not only differ in the precision and reliability of the measurements, but also in the format or formats that are available for downloading, processing and long-term archiving of data. Even though most total stations are sold together with a dedicated piece of software and due to the chaotic situation outlined above, it is normally necessary to use the software provided by the vendor in order to extract data from the device. To put it simply: without a licensed copy of the software, one does not stand a chance of acquiring one’s own survey data. This is a big threat to data preservation, which should clearly be a priority to any archaeological project, regardless of whether it is using proprietary or open source software.

The second side of the problem—the practice—is of course strictly connected to the first and it is related to the practical disadvantages of not having a single and independent framework.
for managing total stations and the data they produce. If multiple programs and workflows are needed for managing all total stations in use, there will be a tendency towards adopting a single vendor’s tools - again leading to data lock-in. Total stations are extremely expensive tools and they require advanced industrial standards, but the same is the case with cars and all cars are permitted to participate in traffic, regardless of how they are built. In contrast, the current situation with total stations (and with other survey tools, not discussed here) reflects a lack of freedom of choice with regards to, for instance, operating systems of workstations and mobile devices. This type of disadvantage can quickly become time consuming if a project or whole company uses more than one model of total station, and this would have a direct and concrete effect on the management and economy of an excavation. It should then be obvious why this lack of freedom should be a concern from the point of view of both commercial or academic archaeologists. These two issues demonstrate clearly where the need for freeing the hardware from any constraint generated by operating system and data format limitations is derived from.

The development of Total Open Station (TOPS) begun in 2008 as a result of frustration by the lack of compatible software in a research team, which was moving to a free operating system (Costa et al. 2013; Costa and Bianconi 2009). The driving forces were a strong necessity to use all the total stations and the data they record indifferently from the operating system used, the fact that no software existed at all for some devices, the desire to make the workflow in the field simpler and faster, adopting a single application to manage all the devices and, in the final instance, to keep a usable long-term archive containing the data itself. In this effort to free the survey instruments from all these constraints, apart from being developed according to an Open Source approach and released under the GNU GPLv3 licence, TOPS forms a step towards “Open hardware”. With the purpose of really liberating the devices and their users, the application can work on many operating systems and it is designed to manage as many as total station models and file formats as possible, with an open structure that allows incremental enhancements.

TOPS is designed for accomplishing basic tasks. It fulfils two main purposes: firstly downloading the raw data recorded with the total station and saving it as is and, in the second instance, exporting that data into common file formats (e.g. CSV, DXF, OGC Simple Features SQL). For completing these tasks in a straightforward way the user interface (UI) aims to be as simple as possible, keeping a good usability level and a user-friendly experience. According to the drive to simplify the use of the application, the UI helps the users in dividing the two parts of the job described above (downloading and exporting). These can also be performed in different instances in time, by processing already downloaded data. The download of data from the total station is the first step. If the device is among the models already supported by TOPS, it can be selected from a dropdown menu and the download is ready to start. If the total station is not already natively supported by TOPS there is a manual mode, in which it is possible to set the connection parameters. Based on user...
feedback, we found out that users frequently ignore these parameters, or have never read the device’s manual. As explained above, the development of documentation and good practices is an essential step that needs to be more widespread.

After the time needed for transferring the data from the device to the computer, this can be inspected as raw data and saved as it is, in a plain text, ASCII format, perhaps the most portable format available. The raw data received from the device is saved by default with a “.tops” extension, but the files can be opened with any basic text editor. It is therefore possible to create a data archive, which can be manipulated later, not based on obscure proprietary data structures but on the most simple and universal data structure: files and directories. It is then possible to open an older data file, and export its data in one of the available formats listed above. Most formats were added upon request by users, providing a tool that is as much flexible as it is simple. The integration of user feedback brings us to another important aspect of the software.

According to open source development good practice the code is hosted in a public repository and a mailing-list has been set up to keep subscribed users updated and, possibly, involving them in the process of development. The same as many other open projects TOPS benefits from users involvement and suggestions in many ways. The feedback from Oxford Archaeology was very important, especially when the application was also tested in a pioneering usage of an open mobile environment (OpenMoko). Similarly, the Gurob Harem Palace Project provided insightful contributions and a solid testing environment for the program, highlighting the obvious problems that every software has. This exchange of information between users and developers is common in open source software projects, and results in mutually shared benefits, leading to a higher quality of software. However, many enhancements are still necessary and numerous new features, related to web and mobile technologies, are planned to be introduced in upcoming releases.

One improvement that has received more attention than others is the ability of TOPS to go beyond a “static” definition of the connection parameters, allowing users to create their own processing workflow just by combining in a single item the connection procedure and the formats of both raw data and exported data. This item will then be used as a preset for all subsequent work, in a way that is similar to other programs specialised in converting data, such as GPSBabel. This functionality is a clear example of how a program can be developed to provide not only an open source implementation of data formats, but also an accelerator for tedious everyday tasks. TOPS aims at being a neutral basis for more elaborate programs, such as the experimental QGIS plug-in that was developed in 2010.

Until recently, open source software was of a rather poor standard in the field of survey, and there are still many improvements to be made, especially with regards to the high quality achieved in the field of GIS tools. Total Open Station is clearly a small component of this next step, but based on our experience we found that its availability is becoming important. It is not surprising then to observe that, at the moment of writing, the program is being included in several GNU/Linux distributions like OpenSUSE, Debian, Fedora and of course ArcheOS. In the near future, it will be possible to plug a total station in any computing device running a free operating system and start working with it immediately.

5. Conclusions

It can be concluded that open source software is very effectively used in the UK, in commercial archaeology and within Egyptian research and rescue archaeology, as well as in other situations.
The manuals produced by Oxford Archaeology represent a powerful demonstration of the potential of open source software and of open methodologies; both have been published online under a creative commons license, and disseminated widely. Whilst initially a slight reluctance to adopt the new work flow was evident, this has now been more or less overcome, and the last two years have seen an increasing number of archaeological fieldwork projects being visualised using open source GIS rather than CAD. In general it has been found that fieldwork project officers and managers are usually more open to adapt whatever technology suits the surveyor best, works fast and efficiently, and outputs the desired results. Given the analysis and query capabilities of GIS it has been possible to convince several archaeologists that on-site GIS can help keep track of the data more efficiently and produce results already during excavation, making the archaeology more accessible to clients and the public. Furthermore, the use of Inkscape by able, but not illustration-specialised archaeologists, speeds up the production of high-quality maps for reports and publications and may lower staff costs. In the same way, the ability to download a total station with any workstation or laptop, in a simple semi-guided procedure, gives more flexibility to surveying both in a commercial and teaching environment. Photogrammetry and other 3D visualisation and analysis methods have recently been adopted in connection with archaeological projects. In this process as many free and open source tools as possible have been employed, such as the visualisation software Paraview. 3D meshes can be imported into GRASS in DXF format and converted to interpolated surfaces using open source software. Nevertheless, it has been found that in other (rare) cases CAD is occasionally still necessary for 3D modelling, in particular for building surveys. However, as the open source GIS packages such as GRASS GIS are becoming increasingly capable, even this use of proprietary software may soon be avoidable. It can be stated that any file formats encountered in our workflow are compatible with any GIS application, and can easily be made compatible with CAD, should this be required. QGIS, GRASS, gvSIG and Total Open Station are able to export ESRI shapefiles as well as DXF files, which can be loaded into CAD, while CAD files can be saved as DXF and imported into GIS. This prevents loss of data and ensures continuity in a system migration context.

The combined use of more than one program to achieve several objectives (data archiving, visualisation and print-quality output) is perhaps not in line with the expectations of those who are used to big “one size fits all” software packages. However, the Unix philosophy of “one tool that does one thing well” is underlying most open source software, and it is not only a key difference between open source and proprietary platforms, but also one of the most appreciated. A common issue on the proprietary side, data exchange, is efficiently avoided in open source software using open and standard file formats: e.g. a map composed in QGIS can be seamlessly edited in Inkscape using the SVG format. In the same way, data can be easily retrieved for other uses, such as advanced spatial and statistical analysis (GRASS GIS).

Since the release of the first two manuals a number of other downloadable resources have been published under the same Creative Commons license. These encompass further documentation on QGIS, PostgreSQL and PostGIS for archaeological use (Robinson 2011a, Robinson 2011b), a guide on 3D Visualisation and Analysis (Hodgkinson 2011b), in addition to guidance on different models of survey equipment (Hodgkinson 2011a). A list can be found in the reference section of this paper. These manuals have also been known to be of use outside commercial archaeology, for instance in university teaching, and their adoption within other industries has been encouraged. These methodologies and manuals have proven successful, archaeological data
maintained within an open source software environment is safer and more sustainable and will be accessible for a longer time.

6. Software Sources
- ArcheOS: http://www.archeos.eu/
- GRASS GIS: http://grass.fbk.eu/
- gvSIG OA Digital Edition: http://oadigital.net/software/gvsigoade
- gvSIG Community Edition: http://www.gvsig.org/web/
- Inkscape: http://inkscape.org/
- PostGIS: http://postgis.refractions.net/
- PostgreSQL: http://www.postgresql.org/
- Paraview: http://www.paraview.org/
- Quantum GIS: http://qgis.org/
- Total Open Station: http://tops.iosa.it/
- Total Open Station QGIS plugin: https://bitbucket.org/iosa/totalopenstation-qgis-plugin

References


Speeding up Georeferencing with Subpixel Accuracy

Gianluca Cantoro
Institute for Mediterranean Studies, Greece

Abstract:
AutoGR-Toolkit version 1.0 is a set of Python (http://python.org/) scripts converted to “.EXE” files with py2exe v0.6.9 (http://www.py2exe.org/). The purpose of the toolkit is to facilitate and speed up the process of georeferencing images with free and open source tools and graphical user-friendly interfaces. It embeds 4 scripts (GGrab, AutoGR-Sift, GeoRef Filtering, GeoTiff Converter) and 2 algorithm libraries (ASift and GDAL) to assist the user in geo-referencing one image on another one according to the specific geographical projection. The use of the software does not require any special skill and it allows the user to go from input to output in few minutes, still keeping the option to customize almost every parameter to improve final accuracy. This paper describes the basic principles and functionality behind each tool in the Toolkit, with some stress tests and results derived from their systematic use. AutoGR-Toolkit is freely available at the URL http://www.ims.forth.gr/AutoGR.

Keywords:
Aerial/Remote Sensing, Georeferencing, SIFT, Free and Open Source

1. Introduction

Georeferencing is nowadays, even among non-specialists, a common practice to warp (scaling, rotating, translating and deskewing) a given image to match a particular size, position and orientation in bi-dimensional real world. Generally the word is used, especially in GIS applications in archaeology, to describe the process of referencing or registering an aerial/satellite raster image to a geographic location in terms of map projections or coordinate systems. Whoever has even a little experience in georeferencing an image knows how unproductive and time consuming this operation may be. At the same time very essential information may be contained in imagery that was produced at a different time or by different source (other camera/lens/filter, other sensor, other medium) or with other lighting conditions. It may therefore be a desire either to combine or to compare this data with that currently available, searching for changes in the features under study, with the highest accuracy possible.

Various GIS tools or specific software are available to rectify rasters into some geographic control framework: ESRI ArcGIS, ERDAS Imagine, PCI Geomatics Geomatica, Global-Mapper, QuantumGIS/GRASS/GDAL, AirPhoto or AirPhoto SE (to name the most well-known). All of them are based on the principle of applying an algorithm to best fit the input with the output (according to the specific input selection and output projection), minimizing the error as much as possible: in a way, they all convert analog information (the human eye/brain which recognizes correspondences in two given pictures) into a digital form (the relative XY coordinate of the input and output raster). Even if this assumption may be extended to the great majority of human-computer interactions, often good (measurable) level of accuracy may be needed for specific purposes but this is not often reachable in a reasonable amount of time.

In the connection between time/accuracy and digital georeferencing output, the free software called AutoGR-Toolkit, developed by the writer at the FORTH-IMS Lab under the auspices of the European Culture 2007-2013 “ArchaeoLandscapes” finds its place. Indeed the proposed software, freely available at www.forth.ims.gr/AutoGR, allows the user to quickly find hundreds of common points in 2 given images.

Corresponding author: gianluca.cantoro@gmail.com
images and geo-reference the one to the other accordingly in matters of seconds.

The software makes use of the principles of the Scale-invariant Feature Transform (or SIFT) algorithm, introduced by David Lowe in 1999\(^2\) to detect and describe local features in images, and now available in different versions or in combination of different algorithms\(^3\) and with improvements\(^4\) or with different libraries (most of them open source).

AutoGR-Toolkit includes also a script that allows the user to grab georeferenced Google images for research purposes using the GDAL library, while respecting the Google disclaimers and policies (abuses are controlled and stopped by Google servers and in no event the author can be liable to the user for issues arising out of the use or inability to use the program, including but not limited to loss of data or data being rendered inaccurate or losses sustained or a failure of the program to operate with any other programs).

In the following paragraphs, details of each tool in the software are provided according to the following scheme: tool description, with technical background information, explanation of the GUI, its functions and expected results; typical applications, where a normal workflow is shown through real examples/tests and case studies; limitations, where the description of unsuccessful applications constitutes the base for future software improvements.

2. **GoogleGrab**

2.1 **Tool Description**

GoogleGrab (or GGrab, Fig. 1) is the first script (of four, in the actual version of the software) accessible through the Windows Graphical User Interface (developed with Python v2.6 and wxFormBuilder\(^5\) v3.2).

This tool helps in defining a draft “background” for image positioning; this is because of the low and not consistent accuracy of Google Maps. Consequently, also the geographical positioning, the pixel spacing and the final precision of the image matched with Google through the use of AutoGR-SIFT may be affected (or compromised) by the aforementioned factors.

Indeed the ASift algorithm (see next paragraph) may be capable of matching traditional vertical or oblique photos with the satellite imagery available at Google Servers, regardless of their different dates, orientation, colour definition and light condition. With this script, potentially every image can be warped to real world coordinates.

---

\(^2\) For the Lowe’s SIFT patent, visit http://www.google.com/patents?id=cleSAAAAMEAJ.

\(^3\) For a combination of ASIFT and SURF, please see Pang et al. 2012.

\(^4\) AutoGR-Toolkit uses the ASIFT (or Affine-SIFT) improvement by Guoshen Yu and Jean-Michel Morel (Yu and Morel 2011; see Appendix).

\(^5\) http://wxformbuilder.org/
GoogleGrab allows the user to save a GoogleMap portion with the area of interest by specifying the top left (North-Western coordinate in the international WGS84 system) and the bottom right corners (South-Eastern coordinate).

A quick screen tutorial assists the user in the retrieval of such information. A “Preview” button in the GoogleGrab window shows a static map of the selected area with linear light-blue edges. If the coordinates are correctly copied/pasted or typed into the form, the preview should depict the area of interest. In the lower part of the window, the user may also specify a different scale (affecting the final pixel size and image resolution) and a name for the output image. Once everything is set, the GDAL library queries the Google server for the specific portion, downloads the area of interest (if the daily limit has not been exceeded and if the area/scale is not too big) in tiles raster format and merges them together. At the same time, a “world file” is stored using the same name as that of the output file to allow its positioning in a GIS system. The same “world file” can be used in the next step (see the following paragraph) to convert SIFT matching points in real world coordinate.

### 2.2 Typical Application

GGrab tool may be successfully used when no other georeferenced imagery is available. Once a single frame or a complete photo sortie is acquired (by military archives or specific aerial archaeology survey) and a first catalogue is produced, the georeferencing process takes place. If the last step is to be accomplished with the use of AutoGR-Toolkit, a Google satellite image may be used for a first draft photo positioning. Next, with the use of other georeferenced images or maps, the final accuracy may be improved and the produced image may be used, for instance, to georeference other images in the same sortie.

### 2.3 Limitations

Specific limitations of use for Google imagery materials are set and are available at Google servers (see above note 6).

Given the already mentioned Google Maps accuracy, a limitation of the proposed tool resides in the impossibility (for the current version) to benefit from the historical images available through Google Earth. This feature, actually under improvement, may reduce the number of non-successfully ASift matching with aerial images of time-period differences.

### 3. AutoGR-SIFT

#### 3.1 Tool Description

This is the core of the Toolkit and it is, for this reason, the most important part of it. As said before, this script “prepares” two input images to be processed by the ASIFT algorithm and converts the output for a GIS environment. The principle of the scale invariant feature transform is quite simple, and also simple is the practical application of it to whatever pair of images selected. Thanks to the pre-

---

6 The retrieval of Google Satellite images (with the specific copyright from Google and from the Satellite program) is to be considered for research and personal purposes only. The user is in any case required to operate according to the latest specific regulations stated by Google, included but not limited to the limitations imposed for top level, storing and manipulation of Google material. For Google Google Maps/Google Earth APIs Terms of Service, please visit https://developers.google.com/maps/terms.

7 Another software created by the same author of this paper, is CataThumb; it allows to easily create an Excel file for cataloguing photo-sorties trough a windows GUI. Its name comes from CATAlogue and THUMBnails (BKMF is an internal code and it was the first name of this software). Given a folder with photos

---

8 The Authors of ASift have put a limit of 3960x3960 pixels for the input images (only PNG format is supported). This is to minimize the required memory allocation and the considerably increasing computing time.

9 “The SIFT features share a number of properties in common with the responses of neurons in inferior temporal (IT) cortex in primate vision” Lowe 1999 p.1150.
compiled version of those algorithms: for any physical object depicted in a digital image, key points\(^\text{10}\) can be extracted to provide a “feature description” of the object itself. This description, taken from a training image, can then be used to locate the same object in another test image (containing many other objects with the one we want to find). It is then essential for the features extracted from the training image to be detectable even under changes in image scale, noise and illumination (Fig. 2).

A good number of publications and cooperative code-writing projects are available (even online) about the SIFT - and variants (see the Appendix)- algorithm and its potentially infinite applications: object detection (Peleato and Jones 2010), robotic mapping and navigation (Se et al. 2001), image stitching (Brown and Lowe 2006), content-based image retrieval (Fraundorfer et al. 2007), 3D modelling (via Structure From Motion), gesture recognition, video tracking and so on.

Nevertheless, none of them includes any processing to be combined with GIS applications (even if image stitching goes very close to it and it starts from almost the same premises).

The AutoGR-SIFT focuses on the GIS image rectification by asking the user to select two images in PNG, JPG or TIFF/GeoTIFF format\(^\text{11}\). Colour space, size, format and dimension of the 2 images may be different and with (potentially) no limitations. Indeed whatever the input images are, they will be scaled and converted in a format suitable for the matching.

Even if the “scale invariant feature recognition” involves reciprocal invariant feature rotation, better results may be achieved by limiting the input images rotation to within a range of 90 degrees\(^\text{12}\). To assist the user in this process, a separate window shows a preview of the two selected images giving the possibility to rotate the second (by 90°, 180° or 270°) to better match the first.

When all the above steps are completed, the matching of the images is undertaken with the Morel and Yu script, and the output saved as structured text files and visual preview of matching points connected by vectors. The list of common points in first and second image is then converted by AutoGR from relative XY pixel coordinate into geographical Easting Northing information according to the input projection\(^\text{13}\).

An automated rectification of the second image will be attempted via the GDAL library

\(^{10}\) Such points usually lie on high-contrast regions of the image, such as object edges.

\(^{11}\) In the next version, a special effort is being put in enriching the number of supported file type, so that input images do not need any pre-processing and can be loaded directly into AutoGR-SIFT.

\(^{12}\) It is also recognized that SIFT is not flip invariant (Zhao 2011)

\(^{13}\) If no world file was available at the beginning of the process, the first image will be positioned at 0,0 coordinate and the second image will be warped on it accordingly.
with predefined settings (Fig. 3): experimental tests with different contexts suggested the parameters to be set as follow: the average residual lower than 10 units with second degree polynomial algorithm, and the number of used points is limited to 100 (if they are more, as usual, otherwise all the points will be used).

3.2 Typical Application

If the accuracy and precision of manual georeferencing of aerial images onto a relatively flat earth surface with user input points are traditionally considered “archaeologically sufficient” (Verhoeven et al. 2012), the image displacements due to the topographic relief usually becomes as important as difficult, especially with sudden steep surfaces and low number of identifiable stable ground control points.

In a recent work (Cantoro and Sarris 2012) the specific software (in combination with CataThumb, see footnote 7) has been successfully applied for the registration of aerial (vertical and oblique, from aerial survey and balloon) images in the case study of a mountainous area of Eastern Crete. Indeed the high potential of AutoGR-Toolkit becomes more visible in mountainous areas, where the local altimetric (sudden) variation requires higher number of ground control point for a correct georeferencing.

Even if AutoGR has been tough for aerial archaeology photo positioning, it can be easily used on lower scale, such as (pseudo) vertical photos for archaeological layers documentation or zoom photos, taken to enrich the details of global images.

3.3 Limitations

For the automatic image matching ASIFT algorithm has been used in this software. It simulates all possible image views and thus is fully affine invariant. ASIFT can handle much higher transition tilts than SIFT (David G Lowe 2004, 91-111), Harris–Affine (Harris and Stephens 1988) a combined corner and edge detector based on the local auto-correlation function is utilised, and it is shown to perform with good consistency on natural imagery (Harris and Stephens 1988, Hessian–Affine (Mikolajczyk and Schmid 2002) and MSER (Matas et al. 2004), but its code complexity is about twice the complexity of SIFT. This results in a slow process (sometime, especially with big raster inputs, the image matching may require up to 10 minutes on a normal desktop computer). Even if some of the matching points may be not correct and filtered out (“false matching”, well known and documented in literature, may appear in specific contexts), this application represents still a considerable speed-up in the traditional manual photo positioning and provides a sub-pixel digital accuracy that has

14 This level of the residual may be considered not acceptable for specific applications, since it may result in 10 meters displacement of the distorted image compared to the first input. For this reason the automatically warped image is presented as a kind of preview and the user will still have the possibility to improve the accuracy by manually deleting incorrect points or by adding new ones in areas outside of the overlapping zone.

15 The Asift algorithm produces hundreds (sometime thousands) of points in less than 3-4 minutes; the computer hardware (essentially processors, RAM and graphic card) may significantly change this value to higher or even lower processing time.

16 The SIFT method has its own wrong match elimination criterion. Nonetheless, it generally leaves behind false matches, even in image pairs that do not correspond to the same scene.
no comparison with the “manual effort”. By providing a list of matching points compatible with the main GIS applications, the script gives an impressive amount of raw data to the researcher for an accurate rectification and for further improvement of error reduction. Nevertheless, the possibility to substitute the ASIFT algorithm with another, such as FAIR-SURF (Pang et al. 2012), may represent a great improvement for the next versions of AutoGR-SIFT. Other limitations, concerning the possibility for the user to choose specific input and output folder, have been fixed for the next version.

4. GeoRef Filtering

AutoGR-SIFT usually produces hundreds of points in few seconds and this third script, GeoRed Filtering, lets the user to reduce this number to a more manageable one (still keeping the whole raw data file). Indeed, according to the specific hardware, the loading more than 200 matching points into the georeferencer utility of ArcGIS and QGis may result in software crashes for lack or bad memory management (Cantoro 2012, 10). This part of AutoGR-Toolkit represented one of the most complex and difficult issue to solve. Indeed the distribution of matching points generated by the Asift algorithm is not homogeneous in the frame area and it doesn’t cover equally every part of it. Instead, the points distribution tends to be denser in high contrast areas and well defined features.

The chosen solution for filtering consists in maximizing the minimum distance between points by selecting one random point of the list and iteratively searching (and selecting) the farther points in the scene. This allows proportional point decimation by preserving their general spatial distribution. Decimation to 100 or 200 points (Fig. 4) should generally be enough for a correct rectification (with second polynomial algorithm) and a good coverage, but since there is a degree of randomness in this filtering (as said, the first point of processing is randomly selected) it may be necessary to repeat the procedure till it meets the specific need or user requirements.

5. GeoTiff Converter

The last script in the Toolkit is a simple Geotiff converter and it allows one to extract the geographical information from any GeoTiff file and save them into a world file using the GDAL library. Some software available on the market are not capable of reading the geographical information from GeoTiff on the fly; instead they need the information to be stored in a “.pngw”, “.pgw”, “.jpgw”, “.tifw” or “.tfw” file. With one click with GeoTiff Converter, the source geotiff file will be converted in regular tiff and jpg images, with world files for both.

6. Conclusions and Perspectives

Even if the AutoGR-Toolkit does not represent something completely new in the field of computer vision or image processing as far as it uses well documented formulas, it surely simplifies the use of complex algorithm mostly intended for other purposes and it introduces a level of accuracy in a reasonably short time in photograph rectification.
AutoGR-Toolkit and the embedded tools (GGrab, AutoGR-SIFT, GeoRef-Filtering and GeoTiff-Converter) are distributed and can be redistributed without charge. The automated installation procedure does not require any special IT skills and also, once the software has been installed by an administrator (high user rights are needed, as for every installation in multi-user systems), it can be easily used by whoever has access to that computer. It does not require any special hardware or software configuration (apart from the free Microsoft Visual C++ 2008 Redistributable Package for computers running Windows Vista) but a good amount of RAM (more than 2 Gb), a multi-processor system (dual or quad core, for instance, better if 64 bit) and a good graphic-card\(^{17}\) may make a difference in processing time (the output itself should be potentially the same in every computer\(^{18}\)).

Several improvements are now in an evaluation process and some others will probably (and hopefully) come from users feedback. They involve different aspects of the Toolkit itself and they have to do with the number of users interested in specific features. For instance, a possible add-on might be to distribute a Linux version of the same software\(^{19}\).

The major improvement which will probably take place shortly involves the chosen algorithm: several alternatives are being considered, which will provide better matching with highly different photos (for instance a satellite with an oblique image of a completely different time-period) and probably in even shorter time. A preference to open source software will be given for the undeniable benefits coming from high customization, user contributions and also to keep on distributing the software in free of charge format. The actual limitation of processing 2 images per run will probably be replaced by a batch processing, to speed up the entire procedure.

This would be of particular value for instance in the georeferencing of photo-sortie from archive or UAV remotely controlled acquisition.

References


\(^{17}\) Tests have been successfully conducted in computer at 32 bit with 1.5 GB RAM and a generic graphic-card and the whole processing times generally does not take more than 5-7 minutes.

\(^{18}\) Efficient recognition is achieved by using local image descriptors sampled at a large number of repeatable locations (Schmid and Mohr 1997, 530). This random sampling may result in different output even on the same machine with the same source photos, if repeated many times.

\(^{19}\) A specific Linux distribution is not yet available, but the software has been successfully tested -as it is now- in a computer running Ubuntu and Wine.
Here is a short list (in alphabetic order) of the most important and interesting projects derived from or inspired by Lowe’s SIFT.

**ANN**


**ASIFT**


**FLANN**


**I-Asift**


**OpenCV**


**OpenSURF**


**SiftGPU**


**Sift++ (now VLfeat)**

Multi+ or Manifold Geophysical Prospection?

Apostolos Sarris
Institute for Mediterranean Studies Foundation for Research and Technology, Greece

Abstract:
Geophysical methods are used as a way of acquiring information related to the subsurface features of archaeological sites. They have been successfully used to map subsurface architectural remains, recognize the limits of settlements and the plan of ancient cities, identify various habitation phases and reconstruct the palaeo-environmental conditions of a region. Even though the current tendency is towards the fast reconnaissance of the archaeological sites through multi-sensor, multi-electrode or multi-antenna systems, the particular approach is deficient in areas that need much more thorough attention. In addition, there are cases where the application of a single method may be inadequate to reveal the underlying archaeological features due to the type and conservation of the features and the surrounding geological and soil context within which they are located. Detailed mapping through the use of multiple geophysical techniques is of increased value since their joint employment can offer a more integrated image of the subsurface. The amalgamation of multiple geophysical techniques can provide complementary evidence of subsurface remains, filling up the shortages that can be produced from a single survey approach and increasing the confidence level of the proposed targets. In order to justify the employment of multiple geophysical techniques and how they address a variety of archaeological questions, the particular paper will draw examples from diverse case studies in relation to the archaeological issues involved and the way of contribution of each method, spanning from shallow depth magnetic and soil resistance mapping to medium depth prospection techniques such as ERT and seismic surveys. Furthermore, it will examine ways of processing various datasets, their correlation and how it may be possible to make the best possible usage of them in terms of the archaeological interpretation. In such a way it will be demonstrated that the manifold geophysical strategy has an outstanding advantage of any other kind of multi-sensor approach.

Keywords:
Geophysical Prospection, Manifold, Multi-sensor, Priniatikos Pyrgos, Szeghalom

1. Introduction
Over the last 20-25 years geophysical techniques have been advanced in terms of sensor technology (faster response, higher sensitivity, less power consumption, storage capacity), mobility of instrumentation (higher portability), semiautomatic navigation (through robotic total stations or GPS units) and the speed of coverage of the sites (fast reconnaissance through the employment of multi-component units and wheel driven, cart based, man pushed, or vehicle pulled systems). The recent emphasis has been given to the multi+sensor component of research through the use of multi-magnetometer systems (e.g. http://www.vallon.de/products.lasso?a=multi-sensor&b=30, http://www.brad.ac.uk/archsci/archprospection/Newsletter1-prt7.pdf), multi-antenna GPR platforms (like those used by VIAS-Geophysical Prospection, Geodesy & Photogrammetry of the University of Vienna & Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology based on a 16 antenna configuration of MALÅ MIRA system: http://vias-geo.univie.ac.at/research-fields/gpr/multi-antenna-gpr/), or multi-electrode soil resistance arrangements (like the system developed by Géocarta, measuring the soil resistance at three different depths (usually from 0 to 50 cm, 0 to 1 and 0 to 2 m): http://hal.archivesouvertes.fr/docs/00/59/25/20/PDF/Archeometric researches and applications 12mai2011.pdf). At the same time, this kind of new generation instrumentation or new design of survey techniques has been also experimented with various hybrid models, combining various diverse prospection...
methods, such as magnetic and soil resistance configurations (like the Geoscan MSP40 mobile platform: http://www.geoscan-research.co.uk/page23.html - Geoscan) or magnetometers and conductivity meters (like the University of Leicester Multi-sensor platform consisting of 6 Caesium magnetometers, EM38 & GPS: http://www.brad.ac.uk/archsci/archprospection/Newsletter1-prt7.pdf) (Campana and Dabas 2011; Dabas et al. 2009; Viberg et al. 2011; Linford et al. 2010; Novo et al. 2010; Gaffney 2008).

The above examples are just indicative and there are more and more companies that are pushing the technology towards the specific direction with the subsequent increase of the cost of instrumentation, confirming the concerns of the anonymous university president: “Why is it that you physicists always require so much expensive equipment? Now the Department of Mathematics requires nothing but money for paper, pencils and waste paper baskets and the Department of Philosophy is better still. It doesn’t even ask for waste paper baskets.” (Barrow & Tipler 1988, 185).

The particular paper is moving between the theoretical issues involved in applying a variety of geophysical techniques for maximizing the perspective of the archaeological sites in terms of the archaeological research questions and the practical implications of such an approach especially when a number of limitations are imposed. In such a way it is proven that a manifold survey strategy is always more productive than a one-dimensional multi-sensor investigation.

2. The Multi+ Versus the “Manifold” Geophysical Universe

In all cases the multi+ (meaning multi sensor) surveys need to comply to specific field conditions. Since the configuration systems are usually bulky they need to operate on a relative flat terrain which needs not to contain frequent obstacles such as field fences, bushes, stones, terraces or trees. As extensive areas of coverage are targeted through multi+ sensor surveys, the geomorphology of the region needs also to be relatively homogeneous, containing either relatively soft soil (for resistivity) without outcrops of bedrock and dense rock distributions, or resistive soils (clay/salt enriched) (for GPR), or low magnetic gradients or noise (geological/anthropogenic) (for magnetometry). Still it is true that they are able to collect a vast amount of high resolution information scanning large areas in a small period of time.

As a consequence, the multi+ sensor surveys produce a vast amount of data that cannot be thoroughly processed in correlation to the specific localized archaeological issues (especially when dealing within a small scale systematic or focused excavation). In addition, extreme confidence is given on the quality of the obtained data without looking for localized singularities that can be related to certain measurements. So, even if multi+ sensor approaches can be consider to address the question “Not see the forest for the trees?” there is still the concern of what is the forest and what is the trees.

In order to understand the issue we shall try to adopt a manifold approach to archaeological prospection. In its simplistic terminology, manifold geophysics correspond to a prospection approach that employs a variety or diversity of methods. In topological terms, a manifold is a topological space that has a local “diffeomorphism” (differences in distances and angles) with respect to the usual Euclidian space. At a small scale, a manifold resembles an Euclidean topology, but in a more global scale a manifold can be much more complicated. It is like we represent a sphere through a mosaic of small tangential surfaces. It is like we imagine a sphere, each small tangential section of which can be represented in a 2D surface, the mosaic of them however can provide a representation of the surface of the sphere. In this way,
Manifold geophysics can adopt methodologies not only to have one level of information (for a particular depth/s and from a specific method), but fuse diverse information from various approaches in order to construct a hyperspace of information that can be used to capture different dimensions of an archaeological site within its environmental settings, as its time envelope unwraps. In our case “more” is with respect to the wealth of information and not with respect to the extensive coverage of the sites.

In order to demonstrate the potential of manifold geophysical approaches we shall draw specific examples from a couple of archaeological sites in Greece and Hungary, where the archaeological questions imposed the use of a diversity of prospection methods.

3. Priniatikos Pyrgos, Mirabello (East Crete, Greece): Shallow and Deep

Geophysical approaches were applied along the course of the Istron Geoarchaeological Project, focusing in the wider area of Priniatikos Pyrgos in the Istron River valley (Fig. 1). Geophysical prospection methods were mobilized to explore the type and distribution of architectural monuments and the intensity and extent of habitation of the particular settlement in various chronological periods, which suggest a diachronic exploitation of the site since the prehistoric times. The archaeological questions concerned not only the diachronic occupation of the site but also the archaeo-environmental issues of the occupation and among others the verification of the existence of an ancient port.

Thus the survey of the area incorporated both shallow depth prospection methods (magnetic, soil resistance and electromagnetic techniques) for mapping the surface anthropogenic relics and deep survey techniques (electrical resistance tomography and seismic methods) for exploring the geological settings of the wider region as a means for reconstructing past environmental conditions that may have influenced the habitation of the region. The above were also complemented with coring and dating (OCL) techniques in correlation to microscopic analyses for the deposition process and reconstruction of the coast line. In addition, a study of the hierarchy of the settlements of the region in different periods was carried out using GIS spatial analysis tools. The particular approach made use of the results of the Vrokastro survey and tried to analyze the dynamics of the settlement patterns and the changing geomorphologic conditions that led to the habitation of the region. Spatial distribution of settlements, clustering of sites, their correlation with the geomorphic attributes of the environment (altitude, slope, aspect, geology, etc), site catchment analysis, least-cost-paths and viewshed analysis were all included among the main processing procedures that were computed. According to them, the geological and hydrographic network, the altitude and distance from the coast line...
and the proximity to cultivated lands played an important factor in the establishment of the settlements in the region (Sarris et al. 2005, 2007; Κουριάτη, 2005).

Strong anomalies indicated by the magnetic data and measurements of magnetic susceptibility verified the existence of a workshop at the centre of the promontory of the Priniatikos Pyrgos. The central section of the promontory pinpointed to the existence of a cluster of kilns, extending over an area of 5-6m in diameter (Fig. 2). The high density distribution of slag fragments at the specific location suggested that at least some of the above facilities may have been used for metal production. Slag fragments, probably derived from iron production activities (corresponding either to the initial smelting reducing process or to subsequent smithing), have been also recovered from other sites in the wider region of Priniatikos Pyrgos (the LM IIIC and later settlement of Vrokastro and the Greek town of Istron located on the promontory of Nisi Pandeleimon, directly east of Priniatikos Pyrgos) suggesting that iron production may have been a significant module of the regional economy in prehistoric and historical times. Massive rubble walls forming corners and rooms were also present along the rocky plateau to the west of the promontory. Multiphase occupation of the site was suggested based on the stratigraphic evidence, partially uncovered due to the gradual erosion of the western scarp. The measurements of the vertical magnetic gradient which were carried out on the top of the west region of the promontory revealed a number of linear features, probably related to the existence of architectural residues. Although most of linear features were aligned in a NW-SE direction, the fuzziness of the magnetic signals suggests that the architectural features may belong to more than one occupation phase.

All of the geophysical data suggested that the locus of habitation and workshop activities was the main headland of the promontory. This was also supported by the magnetic susceptibility measurements: after exhibiting its highest values in the top of the promontory, magnetic susceptibility falls close to background levels (around 35-40emu/gr) to the south of it. In contrast, frequency dependent susceptibility values are kept on relatively high levels (above 10% reaching a maximum of 29.51%). This may imply either that the habitation of the region may have been expanded further to the south of the promontory, or that the specific values may have been caused by the disturbed upper soil horizons. Indeed, the high soil resistance values to the west of Priniatikos Pyrgos and close to the small branch of the Istron River suggested areas of disturbed soil and rubble deposited probable by the repeated cycles of seasonal flooding of the river. Taking in account the direction (towards the sea) of a couple of wide high resistance anomalies (probably correlated to disturbed soil, rubble deposition and other alluvium material carried away from older branches of the Istron River) located towards the SW and SE side of the promontory, it may be also suggested that the older Istron River branches were directed to the sea from both sides of the Priniatikos Pyrgos, leaving probably a small path to the mainland from the SW direction. This is also supported by the sedimentological analyses and OSL datings. If this hypothesis is true, then the settlement...
might have been expanded slightly separate from the workshop area, similarly to what it has been noticed at Gouves (Vallianou, 1997: 333), where in similar river delta settings, the center of ceramic production seems to be separated from the main settlement.

Most of the deep tomographies were conducted in the region between Priniatikos Pyrgos and Nisi Pandeleimon. The extension of the electrical tomography profiles reached the length of almost 200m, achieving a penetration depth of more than 20m at the centre of the profiles. All transects suggested that the superficial cover of the area consists of alluvial/quaternary deposits with resistivity values ranging between 50-150 Ohm-m and variable thickness up to 2-4m. A second layer was also recognized to consist of cohesive sediments with resistivity values between 20-40 Ohm-m. From the depth of approximately 20m and below, a transition zone occurs between sediments and bedrock, probably highly weathered and filled with water, having resistivity values below 10 Ohm-m. Moving further to the south, electrical tomography data suggested an increasing thickness for the upper alluvial deposits (5-6m) which is in agreement to the sloping uphill terrain and a fracture zone that extends below the medium weathered layers. The particular fracture zone seems to shift towards the east and to the south and can be related with a fault, filled with water, which could also explain the corresponding low resistivity zone which appears in the specific region. The above were also visualized through the construction of a pseudo-3D view (horizontal slices of resistivity with increasing depth) created by a 3D inversion algorithm. The particular maps indicate that the fault, which actually has been also suggested by the seismic data, has an orientation NE-SW and it may also be correlated to past heavy landslide episodes that have been reported by the locals originating from the southern hills located to the south of the national road. This faulting could also be responsible for the fracturing nature of the basement which is saturated with fresh water coming from the mainland.

The above were also verified through seismic survey. A general observation of the travel times of the seismic waves suggested that low velocities dominate, indicative of the alluvial deposits that cover the top surface of the area under study. All profiles showed very strong variations of the interface between second and third layer. This was considered to be indicative of the strong velocity anomalies which were caused from the existing inhomogeneity of the bedrock. In an effort to classify the resulted velocity models based on the corresponding formations, the velocity models were distinguished in four different layer velocities as well as four geological formations. Based on this refined categorization (velocity based on geology), four layers were recognized corresponding to specific geological units. According to the particular scheme, the presence of the specific formations in the area was able to be mapped by creating an approximate buffer around the seismic transects. A comparison between the electrical resistivity tomographies and the seismic transects proved that both techniques provided similar results and they also verified the existence of the faulting region in the area between the promontories of Priniatikos Pyrgos and Nisi Pandeleimon (Sarris et al. 2012, Shahrulkh et al. forthcoming).

The 3-D model of the boundary between the sediments (soft and grained materials) (2nd layer) and the bedrock (cohesive conglomerate or weathered limestone) (3rd layer) was constructed in a similar way. Emphasis was given to the area to the west, where a more dense network of the seismic lines was laid, providing a better interpolation for the interface mapping. The depth of the bedrock found to range from 20m to 40m (Fig. 3). It has to be mentioned that the results of both deep electrical resistance tomography and seismic survey verified the detailed categorization of the faults and geological formations that exist in the region. According to the geological data, most of the valley of Istron village, which extends from the national road to the sea,
consists of alluvial/quaternary deposits and fragmentary pockets of limestone sediments. A vertical coastal subsistence of 2-3m that followed the Roman period (Shaw 1990, Raban 1991) submerged a number of coastal facilities in the area that can be seen even today within the shallow waters. The extension of the ancient coast further to the north does not leave a lot of space for natural harbour sites. However, there are a couple of possibilities. One is the estuary of the Istron River that today is observed to flow into the sea from the west of Priniatikos Pyrgos. Although it is possible that its location may have shifted through the past, moving to the west and east of Priniatikos Pyrgos, its estuary may have been used as a protected harbour in antiquity. Another potential location is the areas to the south of Nisi Pandeleimon.

The large promontory of Nisi Pandeleimon, to the east of Priniatikos Pyrgos, may have been separate from the rest of the mainland (as its name “ nisi” suggests). Landslides have been reported by the locals and may have been responsible for connecting Nisi Pandeleimon with the rest of the mainland. The above are also in agreement to the hydro-geological data of the region, which indicate that most of the coastal areas in the vicinity of Priniatikos Pyrgos consist of carstic formations of medium to high permeability and alluvium deposits of variable permeability.

Thus, geophysical data suggested that part of the site (the promontory of Priniatikos Pyrgos) comprised workshop areas during the prehistoric and historical periods, indicating functional similarities to other coastal areas identified on the northern and southern coasts, such as Kommos, Mochlos, Poros and Gouves. The workshop area was probably related to both ceramic production (mainly in the prehistoric periods) and metalworking (mainly in historical periods). Shallow trenching and the analysis of geological cores verified the indications of the geophysical survey, designating among other the limits and type of the settlement and the shifted location of the Istron River which is probably responsible for the deep colluvial deposition noticed in the wider coastal area. A faulting region in the area between the promontories of Priniatikos Pyrgos and Nisi Pandeleimon and the weathered limestone bedrock may also designate the location of an ancient harbour in the specific region.

4. Hungary: Macro and Micro Investigations – the Forest and the Tree / the Primitive and the Processed / the Seen and the Unseen

With the contribution of the USA-NSF (U.S.-Hungarian-Greek Collaborative International Research Experience for Students on Origins and Development of Prehistoric European Villages) and Wenner-Gren Foundation (International Collaborative Research Grant, “Early Village Social Dynamics: Prehistoric Settlement Nucleation On The Great Hungarian Plain”), it became possible to continue the geophysical investigations at Sceghalom-Kovácsalom at E. Hungary. The campaign was carried out under the auspices of KRAP (The Koros Regional Archaeological Project) which since 2002 has been investigating various Neolithic and Early Bronze settlements of central eastern Hungary. The goal of the geophysical research, which was accompanied...
by further modules of KRAP project such as surface surveying, test excavations, coring for chemical analyses, topographic mapping, a.o., was to contribute in the understanding of the formation of the tells and the evolution of the farming societies in SE Europe (Yerkes 2010).

Extensive prospection research was contacted at Szeghalom which is located at the agricultural area to the north of the Sebes Körös River and SW of the village of Szeghalom. The site is easily accessible through a modern levee road above the Sebes Körös (Fig. 4). A drainage ditch is running parallel to the levee to the south of the site. Szeghalom is a representative Tisza settlement with some intrusive Bronze Age finds (Ecsedy, et al. 1982: 161). The predominant Late Neolithic character of the site was also confirmed by a recent reconnaissance of the site that was carried out in 2008 by Salisbury and Pultz and who observed large concentrations of diagnostic sherds and daub (Salisbury 2010). Salisbury carried out geochemical, magnetic susceptibility measurements and soil texture analysis to characterize the use of space and the type of activities in the area focusing on two specific locations (Loci 1 and 2) to the south of the tell.

The geophysical prospection surveys that were initiated in 2010 employed extensive magnetic survey and GPR. Further modules of the research connected to the geophysical prospection campaign involved the use of aerial and satellite images and the experimental spectroradiometric measurements above suggested geophysical targets. The goal of the geophysical research was to map the subsurface architectural relics on the tell and in the area extending around the tell to provide additional information regarding the extent of the sites and the evolution of habitation patterns of the region. More than 250,000 square meters have been covered through the use of the magnetic techniques and about 26,000 square meters were scanned through the use of the GPR. In 2012 and ERT survey was also carried out in order to conclude on the stratigraphy of the tell and the direction of the paleochannels.

The survey (by almost all methods) on the tell suffered from problems originating from the multiple occupation phases existing in the region. The same holds for the region to the south of the tell. Furthermore, plowing activities created increased noise levels at different sections of the site. The GPR measurements were taken with extreme difficulty due to the muddy environment of the tell and its surroundings. The GPR slices with increasing depth indicated that as we move outwards from the limits of the tell, we are falling on water saturated soil and the radar signals are heavily attenuated. The rest of the area on the tell produced well registered reflections, most of which appear at the top of the tell. As it was the case of the magnetic anomalies, and as it was expected due to the past excavations in the specific area, the produced reflections were noisy and fuzzy, and thus only a small number of possible linear features was possible to be distinguished. A number of linear features that belong to plow marks or tractor tire marks (sometimes reaching the depth of even 30-40cm below the surface) were also registered in the measurements (Fig. 5).

Thus, most useful information came from the magnetic survey even if it is difficult to recognize isolated anomalies. According
to the magnetic survey the tell seems to have dimensions of about 120x75m covering a total area of 9,000 square meters. According to the classification made by Kalicz and Raczky (1987:16) for the Tisza period tells (between 1-4 ha), it seems that Sceghalom tell belongs to the lower size category of tells. Contrary to the cases of Vésztő 20, Kórösladány 14 and Vésztő-Mágor tell, magnetic survey did not provide substantial evidence for the existence of ditches around Sceghalom tell, confirming the older plan of the site that depicts also the location of the past excavation trenches. What seems much more probable is that the tell was at least partially fortified with palisades and a system of meanders was running around it forming a natural defensive ditch. The GPR survey has given some partial evidence on the identification of the traces of the past excavation trenches.

Magnetic data were of particular success in identifying the residues of structural remains around the settlement, most of which have been seen to be intentional burnt (as the 2012 excavations have shown). The flat settlement seems to expand in all directions around the tell. The built environment seems to develop after leaving a buffer zone of about 100-200m of un-built area from the tell, making a good exploitation of the surrounding wetland environment. Until now there is no indication of an enclosure around the settlement. Still, differentiations exist within the built environment of the area, distinguishing two sections, one densely settled and another one of more dispersed character. Still it is not known if the particular flat settlement sections were created as a result of a gradual population movement or due to pressure for more land exploitation.

To the north, northeast and southeast, a dispersed type of settlement consisting mainly of Tisza type farmsteads is suggested (Fig. 6). More than 40 such structures of dimensions around 17-25m x 8m have been identified. These houses of oblong ground plans have similar dimensions and orientation and some of them can be considered as thermal targets, pinpointing to intense heating/crafting activities in their interior. Most of the magnetic features can be correlated to fired daub walls, postholes, or other thermal targets (e.g. kilns and fire hearths). Similar were the results of the magnetic susceptibility. On the other hand, the GPR had limited success in outline the structures and no success in detecting smaller features like pits.

It seems that the GPR signals were influenced by the small terrain fluctuations and the plow marks as well as the increased humidity of the soil and the filled with water depressions.
of the terrain that they were located at various sections of the site. The effects of water saturation in soils to the GPR signals became evident in the survey of a number of grids, especially those extending to the low elevation depressions of the site. The large attenuation of the electromagnetic waves within the water saturated soils obscured completely features that were registered well in the magnetic survey. If the prospection of the site was going to be relied only to the GPR surveying, more than half of the structural remains would have been revealed (Fig. 7).

5. Final Remarks

Both case studies from Greece and Hungary indicate that the use of a single method would not be sufficient for addressing all the archaeological questions regarding the settlements. The manifold geophysical approach employing a variety of techniques was the only way to explore the shallow and the deep strata of Priniatikos Pyrgos, investigate the environmental settings of the site, and conclude on the past coast line reconstruction. Similar was the case of Szeghalom, where the conditions of the ground were imposing specific limitations on the type of the prospection method that could yield the best results.

Past efforts of integrating various geophysical techniques have been carried out with success (Kvamme 2006; Keay et al. 2009; Ernenwein, 2009). However, the recent trend of the employment of multi-sensor instrumentation has shifted the focus of the prospection survey towards the fast reconnaissance of the sites using one module of survey, deviating severely from the actual multidimensional archaeological questions. Despite the recent developments towards a multi+ sensor geophysical technology, most of the archaeological sites cannot be approached through the particular instrumentation (since it is bulky and difficult to employ in rocky, bushy and sloping terrains). Furthermore, the intrinsic risk of relying to a single technique may hinder the manifold revelation of the various dimensions of a cultural space.

Acknowledgments

With the contribution of the USA-NSF (U.S.-Hungarian-Greek Collaborative International Research Experience for Students on Origins and Development of Prehistoric European Villages) and Wenner-Gren Foundation (International Collaborative Research Grant, “Early Village Social Dynamics: Prehistoric Settlement Nucleation On The Great Hungarian Plain”), it became possible to continue the geophysical investigations at Szeghalom-Kovácszalom and carry out some further experimental work at Véstő-Mágor (Vésztő 15). KRAP is indebted to both agencies for their support. Finally, funding for making the presentation at the conference was provided by the Culture 2007-2013 Archaeolandsapes Europe project.

References


Impact Assessment: The BREBEMI Project (Italy).”


Shahrukh, M., P. Soupios, N. Papadopoulos, and A. Sarris. in press. “Medium depth geophysical investigations at the Istron archaeological site, eastern Crete, Greece using 3D Refraction Seismic and Geoelectrical Tomography.” Journal of Geophysics and Engineering.


Managing Data from Multiple Sensors in an Interdisciplinary Research Cruise

Øyvind Ødegård, Martin Ludvigsen, Geir Johnsen, Asgeir J. Sørensen, Stefan Ekehaug and Fredrik Dukan
Norwegian University of Science and Technology, Norway

Mark Moline
University of Delaware, USA

Abstract:
The Applied Underwater Robotics Laboratory (AUR-Lab) at the Norwegian University of Science and Technology (NTNU) is an inter-faculty arena for multi- and interdisciplinary research and education on underwater robotics, marine acoustics, marine biology and marine archaeology. The main scientific focus of AUR-Lab is development of technology for ocean observation comprising guidance, navigation and control of underwater robots as instrument carrying platforms and corresponding use of sensors to identify, map and monitor sea surface, water column and seafloor. In December 2011 AUR-Lab conducted a research cruise to test its ability to run integrated operations deploying ROVs and AUVs as remote sensing platforms for baseline surveys. A key element in the operations was data and information interdependency, with data processing and interpretation in near real time. Even though the data processing pipeline can be further optimized, the concept was proven. The cruise was considered a success, and scientific results were obtained together with valuable operational experience.

Keywords:
AUR-Lab, AUV, ROV, SSS, MBES, Underwater Marine Archaeology and Bio-Geo-Chemistry

1. Introduction

Europe’s ocean space is under increasing pressure from human activities occupying coast lines and shelf areas. One strategy for facing growing industrial activities is to move coastal activities further offshore. With an estimated 20 % increase in gross value by 2020, EU’s marine and maritime economies are seen as an “untapped potential” that will help Europe to recover from the current economic crisis (European Commission 2012). Acknowledging the challenges this massive expansion into the marine domain entails, there is currently a strong focus on research on marine spatial planning and “building up knowledge relevant for integrated and ecosystem based management should therefore be emphasized and facilitated” (JPI Oceans 2012a).

The effects of this development can already be seen along the Norwegian coastline. There are plans for offshore wind farms, increased exploitation of hydrocarbons in deeper waters and a steady growth in fisheries, aquaculture and shipping activities. A safe and sustainable management of the ocean space in the future requires a strong commitment to research for developing adequate methods and technologies, and educating the next generation of marine scientists and technologists. This will ensure understanding of the marine ecosystem and to provide tools for better identification, mapping and monitoring of areas of interest. It is important that this effort is conducted across the traditional marine and maritime sectors, ensuring a common understanding among all stakeholders regarding the challenges ahead.

According to the Norwegian Cultural Heritage Act the state has ownership to wrecks with belonging gear and cargo, or parts of such objects, if they are more than 100 years old. Anyone planning to initiate measures, which may affect such wrecks, must send their plans...
to cultural heritage authorities for a statement on the potential for conflict. To provide such a statement, the authorities often need to do a marine archaeological survey of the areas in question. Usually such surveys are conducted by diving marine archaeologists using standard scuba equipment, and never deeper than 30 meters. For larger or deeper areas, where diving operations are impossible or less effective, use of remote sensing technologies is necessary.

1.1 Applied Underwater Robotics Laboratory (AUR-Lab)

For the last 10 years NTNU’s strategic program Marine Coastal Development (MCD), has been one of six main strategic research fields at the university. In 2009 the Applied Underwater Robotics Laboratory (AUR-Lab) was established as an inter-faculty research infrastructure for MCD partners. AUR-Lab represents a new arena for multi- and interdisciplinary research and education on marine underwater robotics, marine acoustics, marine biology and marine archaeology. The main scientific focus of AUR-Lab is development of technology for ocean observation comprising guidance, navigation and control of underwater robots and corresponding use of sensors to identify, map and monitor sea surface, water column and seafloor. Further on, the application of these technologies on different science areas is also a major goal for AUR-Lab. Research groups on technology, biology and archaeology work across disciplines to develop new knowledge, solutions and methods for marine operations and research in the ocean space. A close collaboration with both industry partners and different government/management agencies enables AUR-Lab to approach directly the research gaps defined by EU (JPI Oceans 2012b). AUR-Lab currently manages an infrastructure comprising a research vessel (RV Gunnerus) equipped for launch and recovery of different underwater robots; two Remotely Operated Vehicles (ROV) and an Autonomous Underwater Vehicle (AUV).

Prior to planned development projects that could have an effect on the marine environment a common management strategy is to conduct baseline surveys to gather reference data for environmental impact assessments. Typical events trigging baseline surveys are oil and gas development projects, general environmental mapping of natural resources and establishment of aquaculture or construction of wind farms.

Aspiring to develop methods and technologies that are applicable for multiple end users, AUR-Lab conducted a baseline survey mapping objects or areas of biological and/or archaeological interest. This would enable technological and operational capabilities to be tested within a very realistic framework. An area was chosen near the island Frøya on the outer coast of central Norway that had not been surveyed before.

Integrated Operations is a term coined by the Petroleum Industry referring to work processes and ways of performing tasks facilitated by new information and communication technology used in multi-
discipline cooperation. Applied to AUR-Lab research cruises, the idea was to reduce redundant repetitions of tasks at different work stations, thereby streamlining operations by optimizing the utilization of both technological and human resources. For this research cruise AUR-Lab would first map the seabed with Multi Beam Echo Sounder (MBES) and AUV in overlapping sequences. AUV survey lines would depend on the bathymetry gathered by the MBES. Then possible objects or areas of interest would be identified during processing and analysing of the combined datasets for biological and archaeological investigations by ROV or SCUBA divers. The overall goal of the cruise was to test if a baseline survey could be effectively conducted utilising multiple sensors and multiple platforms within an Integrated Operations framework. If successful it would present a new model for cost effective baseline surveys while retaining high scientific quality.

A list of criteria for measuring success was formulated for the whole cruise and for each sub-operation. This paper will give a short introduction of the technologies and concepts involved in the cruise, a brief summary of the cruise followed by a discussion of the benefits and challenges this approach to integrated operations poses for marine archaeology.

1.2 Study Area

Dragnessvaet on the north side of the island Frøya is a relatively open bay surrounded by small islets and low rocks giving some meagre protection against heavy seas from the east and west. The area could be said to have typical characteristics of the outer Norwegian coastline with a rocky landscape, exposed to heavy weather and a resulting coarse-grained sea-bed substrate. Frøya has seen much maritime activity and traffic ever since the Mesolithic period – especially related to rich fish resources. With no potential for prehistoric landscapes in this region, and since the area is off the main fairway, the expected potential for underwater cultural heritage is limited to wrecks or other traces of local small scale seafaring.

2. Methodology

By and large the technologies and equipment required to conduct a baseline survey exists within AUR-Lab’s infrastructure. Unless stated otherwise the equipment mentioned is owned by NTNU.

2.1 Research Vessel

The RV Gunnerus is 31 meter long with a personnel capacity of 25. It has a large work deck, a well-equipped workshop, wet and dry labs making it very well suited for AUR-Lab research activities. A Kongsberg Simrad Dynamic Positioning system (SDP-11) combined with Kongsberg Seatex Differential Positioning Sensor (DPS-232), Kongsberg High Precision Acoustic Positioning system (HiPAP 500), and a Kongsberg Seatex Motion Reference Unit (MRU-5) integrated in a Kongsberg Seaphath 300 system, ensures high precision references for all kinds of sensors, instruments and underwater control operations. An important part of the cruise was to test and qualify the vessels capability to launch, recover and control all relevant sensor platforms in an effective manner – running parallel operations if required.

2.2 Multi Beam Echo Sounder (MBES)

The ship is fitted with a hull mounted Kongsberg Maritime EM3002 single head MBES with dynamically focused beams. The EM3002 can operate in depths from 0.5 to 200 meters, with a swath width of 10 x depth (max 200m). It is well suited for high resolution seabed mapping in shallow waters. The sonar head is fixed to the hull of the ship, and is therefore limited to the operational radius of the vessel. Kongsberg’s Seafloor Information System software (SIS) was used as interface and for real time processing.
The EM3002 produces a dual data output. The primary data output is the bottom detections which gives depths that can be used to create bathymetry. Bathymetric data is collected as x,y,z points, where a function of time and angle gives the positions of seabed echoes relative to the echo sounder. The spatial resolution of each point depends on several variables, such as array aperture, vessel speed and depth.

In addition to bathymetric point data, the MBES also produces a seabed image based on the acoustic backscatter from the seabed. This image is somewhat similar to a Side Scan Sonar (SSS) image, but with some notable differences. As an SSS usually is towed close to the seabed to obtain a small incidence angle, the main purpose is to detect shadows of features. An MBES produced seabed image is based on data collected at a much larger angle of incidence, and therefore better suited for seabed classification (Hammerstad 2000). For this particular survey the main purpose of the MBES was to create a terrain model for planning the AUV survey lines for optimum and optimal coverage. The MBES backscatter was used to complement the SSS data when selecting targets. The IVS3D Fledermaus Pro software suite was used to process and analyse MBES data (both bathymetry and acoustical backscatter).

2.3 Autonomous Underwater Vehicle (AUV)

An AUV is a robot that can operate on its own without real-time control from a surface operator. Autonomous robotics is a very rapidly evolving field of research, and AUVs are being applied to more and more advanced operations. An AUV can be programmed to swim in a predefined pattern and at predefined depths. Depending on its sensor setup and processing capabilities it can also adjust on the fly to certain sensor inputs, i.e. it can measure its distance to the seafloor with a Doppler Velocity Log DVL and maintain a constant flying altitude following the bathymetry of the seafloor (Moline et al. 2005).

After finishing its mission, the AUV returns to a predefined area in the proximity of the launch vessel, surfaces and beacons its position by WiFi and strobe light. After recovery the data from the mission is downloaded and processed on board the research vessel. After recharging its batteries it is ready for a new mission.

The AUV used on the cruise was the Remote Environmental Monitoring UnitS (REMUS) “Boomerang II” owned by California Polytechnic State University, which is similar to an AUV that AUR-Lab has purchased. The REMUS 100 model is a modular propeller driven
AUV developed by the Oceanographic Systems Laboratory at the Woods Hole Oceanographic Institution, and is now produced by Hydroid, Inc., a subsidiary of Kongsberg Maritime. It weighs 37 kg, is 160 cm long and has a 19 cm diameter. Four Li-ion batteries give the AUV a range of approx. 80 km at a speed of 3 knots. In addition to a flux gate compass, a system of four Long Base Line (LBL) two-way communication transducers were deployed around the survey area helping the AUV to navigate while submerged.

2.4 Side Scanning Sonar (SSS)

A Remus 100 AUV (Kongsberg) was equipped with a Marine Sonics 900 kHz SSS that produces seabed images based on acoustic backscatter data. As for the MBES, the backscatter is measured in terms of echo strength. Since the AUV can obtain a flight path close to the seabed (in this case 3 m altitude) and the SSS is directed sideways, the backscatter strength describes the grazing angle more than hardness or absorption characteristics of the seabed. The low flying altitude following the seabed produces SSS datasets that describe the seabed in a very consistent way, easing interpretation and comparison with other datasets. A low grazing incidence produces acoustic shadows much larger than the features that are being ensonified, which are very useful for identifying objects. The combination of bright reflections and black shadows gives a very photo-like monochromatic representation of the seabed and for better and worse eases intuitive interpretations (for SSS details, see Blondel 2009).

Considering the potential for underwater cultural heritage in the study area, one must take into account the very mechanical erosive conditions. Few structural elements of wooden hulls can be expected to remain, and a typical wreck site would have a very low profile with objects scattered over a wide area by the winter storms. In this research cruise the AUV was used as a platform for gathering SSS data for detailed mapping of features of archaeological or biological interest. A special focus was given to the ability to determine presence of kelp forest using SSS data. Chesapeake SonarWiz 5 software was used to process and analyse the SSS data. The SSS data was then exported as Geo-TIFFs to be draped on to the bathymetric model in Fledermaus for further analysis.

2.5 Remotely Operated Vehicle (ROV)

An ROV is connected to a surface control unit by a tether, often referred to as an umbilical. The typical ROV is equipped with a set of thrusters configured to allow manoeuvrability in all directions. All ROVs have as a minimum lights and a video camera, but can be fitted with all kinds of sensors and instruments depending on its payload capabilities.

AUR-Lab currently has two ROVs; one work class ROV custom built for archaeology, and one smaller inspection class ROV. On this survey MINERVA, the inspection class Subfighter 7500 was deployed and operated from a 15 ft. container on the aft deck of RV Gunnerus. MINERVA has a maximum operating depth of 700 meters, and a payload capacity of approximately 20 kg. It is manoeuvrable in up to 2 knots current speed, and is well suited for inspection of selected targets and conducting video/photo surveys along predefined lines. MINERVA is equipped with several camera systems, and can be fitted with additional sensors if required. High accuracy positioning is maintained by an automatic Dynamic Positioning (DP) system developed by AUR-Lab for MINERVA (Sørensen et al. in press). This system utilises a Doppler Velocity Logger (DVL), a Motion Reference Unit (MRU) and a transducer that measures a relative position to the hydro acoustic positioning system HiPAP 500 system installed on RV Gunnerus to navigate and track the position of the ROV. Such high accuracy positioning enables the ROV to be an adequate platform for photo mosaic and photogrammetry cameras. To test and prove the ability of MINERVA’s DP system
to facilitate the acquisition of photomosaic and photogrammetry imagery was one of the goals of the research cruise. MINERVA cameras would inspect targets selected from the gathered MBES and SSS data to verify or reject them as objects of interest.

Other relevant sensors like sub bottom profilers or magnetometers were not available for this research cruise and extending the range of large sensors would also have exceed the payload capacity of the MINERVA ROV.

2.6 Data Management

The concept of Integrated Operations has data and information management as a key element. A whole range of both commercial and AUR-Lab developed software is needed to operate all the different platforms and sensors. Streamlined operations depend on a smooth transfer of data and information between workstations – it should for instance be easy to pull navigation data for an AUV-mission from a processed MBES dataset with a minimum of delay due to format transformations. A solution based on a common server with a functional level of data interoperability would be desirable. Since most of the relevant data and information would have spatial properties, ESRI's ArcGIS was selected for data management. In addition ArcGIS was to be used for the project's final data storage (including metadata, logs etc.) and report map production.

The idea was that all spatial data should be made available to different researchers for independent analysis, i.e. archaeologists and biologists would have privileges to the same geo-database and use it as a common work space. Also the ROV-operators should be able to pull navigational data for programming the ROV from archaeological or biological target layers in the geo-database. A Mac Pro (two 2.26GHz Quad-Core Intel Xeon 5500, 16 Gb RAM) was set up with VMware fusion software to virtualize two Windows 7 (64-bit), and one Windows XP (32-bit) environments to facilitate different software requirements. IVS Fledermaus has the ability to interface directly with the ArcGIS geo-database, and was meant to be the main analysing tool since it is very well suited for visualising 2D- and 3D-data simultaneously. Fledermaus and ArcGIS could only (in December 2011) communicate in a 32-bit Windows environment, and a video recording extension (Mapix) for ArcGIS would only work in Windows XP. It was planned that connections between Fledermaus and a ArcGIS geo-database was to be set up and served in

Figure 3. A simplified scheme showing the different data sources and formats with corresponding software.
Windows XP (32-bit), while more process intensive operations (e.g. rendering large 3-D or raster datasets) was to be done in Windows 7 (64-bit) environments.

3. Results

At first a MBES survey was conducted to cover the area of interest and to gather data for a bathymetric model. The MBES data was cleaned manually using IVS Fledermaus as it was suspected that using a Combined Uncertainty and Bathymetry Estimator (CUBE) algorithm could possibly interpret cultural heritage as unwanted noise. Dispersed and low profile features will produce ephemeral acoustic signatures, and are not very likely to be detected by MBES alone.

The MBES data from each survey line overlapped by minimum 15-20 % to ensure high quality coverage for the whole area.

The MBES data was then analysed by both archaeologists and biologists to detect possible Objects Of Interest (OOI) such as archaeological objects and to pick areas with potential for presence of kelp forest and corresponding soft and hard bottom substrates. After some experimentation with grid settings and different shading options in Fledermaus, it was found that the best way for analysing MBES-data was to view it as a point cloud showing every single ping. Survey lines for the AUV were then defined based on the bathymetric model, to give good coverage of the identified targets and areas.

This part of the survey was conducted in two phases. In phase one, approximately half the survey area was covered with MBES. In phase two the MBES survey continued; while the AUV surveyed the area covered during phase one – the bathymetric model having been generated immediately after the last survey line was finished. This was an essential part of the Integrated Operations concept, and one of the success criteria for the cruise.

After having downloaded the high resolution SSS data, the data were processed and analysed using Chesapeake SonarWiz 5 software. Then they were exported as Geo-Tiffs for use in Fledermaus and ArcGIS. The compilation of data sets was done in IVS Fledermaus giving the researchers the opportunity to analyse data in both 2D and 3D views. Having the possibility to turn layers on and off, and also to drape seabed images on top of the bathymetric model, a more complementary impression of the seabed was attained.

The underwater terrain of the surveyed area was hilly with rocks and sand, and depths varied from 10 to approximately 70 meters. A total area of 1.34 square kilometers was covered during the cruise, and 11 targets identified for closer inspection. Survey lines for ROV inspection was planned for each target, and the DP system proved to function well within the margin of error defined as a criterion for success for all targets.

All targets identified by archaeologists were rejected as cultural heritage after visual inspection by ROV. All areas suspected to have

Figure 4. MBES and SSS coverage.
presence of kelp forest were confirmed by ROV video inspection.

3.1 Data Management

Different work stations and software suits operated with different datums and projections. While it could not be concluded that it was impossible, the effort required to set up SIS, Eiva Navipac, Olex and ArcGIS to exchange survey lines far transcended the time and capabilities at disposal during the cruise. After a few fruitless attempts a less sophisticated solution based on third-party software for transforming coordinates and plotting lines manually at each work station was adopted. The extent of this problem was not sufficiently understood during the planning of the survey, and it was a consistent source of delay throughout the cruise.

To have multiple instances of ArcGIS and Fledermaus running on different OS share a geodatabase in a virtualised environment turned out to be problematic. It is not clear what exactly the problem was, but after repeated blue screens and an increasingly slower dataflow, this concept had to be rejected in favour of a folder based file structure.

AUR-Lab was able to establish a bathymetry model of the area based on MBES data. The area of interest was also covered using SSS from an AUV. AUR-Lab demonstrated the ability to investigate and document targets by ROV based video survey. During the AUV operation measurements of seawater quality (temperature, salinity, coloured dissolved organic matter and turbidity), and current speed and direction (ADCP measurements) was collected.

During this cruise a DP system for ROV developed at AUR-Lab was tested. AUR-Lab launched release 1.0 during this cruise and the cruise will hence serve as a SAT. The ROV was capable of manoeuvring in a lawn-mower pattern with a constant seabed altitude and with a cross track error less than 1 m.

4. Discussion

Use of remote sensing under water is not unknown in marine archaeology. Side Scan Sonars (SSS), magnetometers, Multi Beam Echo Sounders (MBES) and Sub Bottom Profilers (SBP) have long been used in archaeological investigations (Bates et al. 2007; Quinn et al. 2005; Wunderlich et al. 2005). Considering the very diverse nature of underwater cultural heritage, a great deal of attention has been paid to the different sensors ability to detect objects of possible archaeological interest. A much recurring topic in those discussions has been, and still is – resolution and the ability to discriminate cultural objects from natural features (Bates et al. 2011). The classical trade-off is between coverage and resolution, mostly due to the acoustic properties of water. Higher frequencies produce a higher resolution, but at the same time a shorter range. Synthetic Aperture Sonar technology seems to produce data that to some degree transcends this limitation (Hansen and Hagen 2009), but so far this technology has not been applied to archaeological purposes.

"Such amounts of high-resolution data from multiple sensors will reach the limit of human based interpretation. There is a need
for new multisensory data fusion tools to help in the processing and exploitation of modern surveys” (Alain 2012).

One of the objectives for the archaeological participation in AUR-Lab is to develop better and more effective methods for conducting underwater cultural heritage surveys. Integrating data from multiple sensors on multiple platforms could enable researchers from different disciplines representing different interests to conduct common surveys – all benefitting from extended data gathering. This concept is not entirely new (Ludvigsen and Soreide 2006), but our survey is the first fully integrated survey involving several disciplines utilising multiple sensors and platforms in Norway.

A long path length between sensor and seabed produces data with a large swath width and low spatial resolution. This makes it harder to distinguish between cultural heritage and natural features (rocks etc.), and could also make the survey fail to detect more indistinct – and perhaps older, traces on the seabed. One way to rectify this problem is to place the sensors closer to the seabed. A multilevel approach to detection of underwater cultural heritage can be achieved with the appliance of underwater robots. Using submerged platforms (e.g. AUV or ROV) for deploying a range of sensors, a shorter range will produce smaller footprints and thereby a higher resolution than for the traditional hull mounted sensors.

Complementary sensor data can narrow down the list of possible targets, and increase the probability of actual archaeological relevance for the remaining targets. If archaeologists can collaborate with other research fields, industry, management agencies and decision makers with respect to finding methods with multidisciplinary benefits, the resulting increase in survey efficiency and quality should be obvious. However, such solutions entail a high level of complexity, and will thus require a high level of competency among the participating disciplines and increased investing in research and method development.

5. Conclusions

An increased exploitation of the ocean space presents new challenges for the underwater cultural heritage management. A growing demand for integrating archaeological surveys into existing survey regimes calls for innovative approaches to adapting technologies and methods in integrated operations.

In December 2011 NTNU AUR-Lab conducted its first integrated operations research cruise. With an ambitious plan for scope of work, the main objective was to test the ability to run integrated operations were data and information interdependency would be an essential element.

A total of 11 targets were identified and investigated analysing combined data from multiple sensors deployed on both surface and subsea platforms. Ground truthing was conducted on selected targets by ROV video surveys.

The cruise was considered a success. Scientific results were obtained together with valuable operational experience. Despite some obstacles the planned operations were executed, and to some degree the planned workflow, albeit altered, could be followed. Even though the data processing pipeline was far too slow, the concept was proven – and improved implementation will no doubt dramatically increase overall efficiency for such operations in the future.

AUR-Lab is currently working towards expanding its infrastructure with a subsea test centre including an AUV docking station, ROV for arctic operations and time series stations for biological and archaeological research.
Acknowledgements

The authors are grateful to all participants on this AUR-Lab cruise, especially the crew of RV Gunnerus for their invaluable help and positive attitude. The authors also wish to thank the anonymous reviewers for constructive advice that improved the quality of the paper, particularly reviewer B’s comments were very helpful.

References


Towards Detection of Archaeological Objects in High-Resolution Remotely Sensed Images: the Silvretta Case Study

Karsten Lambers
University of Bamberg, Germany

Igor Zingman
University of Konstanz, Germany

Abstract:
We report on recent research undertaken in the framework of the Silvretta Archaeological Project, in which we are developing methods to detect certain types of archaeological ruins in remotely sensed images in order to assist archaeological survey. Our approach aims at assessing the probability of the presence of objects of our interest based on geometric cues that can be automatically detected in the satellite and aerial images that we use. We describe our methodology and the first integral step, constituting a new approach to texture segmentation that we developed to reduce the rate of false detections.

Keywords:
Archaeological Survey, Remote Sensing, Texture Segmentation, Texture Contrast

1. Introduction

In his keynote lecture during the opening session of the Southampton CAA conference, Jeremy Huggett identified digital image analysis as one of the great challenges of digital archaeology for the next few years (Huggett, this volume). We agree because the amount, variety, and availability of digital images as well as their actual or potential use in archaeology, has increased dramatically since the introduction of digital optical sensors in earth observation remote sensing in the 1970s (Giardino 2011). With regard to image analysis, it seems however that archaeology lags behind other disciplines (De Laet and Lambers 2009, Lasaponara and Masini 2012, Cowley 2012). For example, in such diverse fields as earth observation, surveillance, transportation, medical imaging and social media, digital image analysis is today not only used routinely to correct or enhance images, but also to analyse their content, which often includes object detection using computer vision approaches. Automated visual analysis has substantially advanced in recent years, now allowing for a variety of targets to be automatically detected, such as buildings, roads, people, or various patterns in medical images. Moreover, remarkably successful algorithms and technologies have been developed for face detection and for object detection for autonomous car navigation (Szeliski 2010). In spite of remaining limitations, these examples demonstrate the potential of automated object detection across different disciplines.

In contrast to the aforementioned sciences, in archaeology, the interpretation of remotely sensed images, including the detection of archaeological sites and objects, is in most cases still the domain of archaeologists visually inspecting and interpreting the images. Just a few recent studies go beyond that level by attempting to correlate spectral image properties and archaeological site distribution (e.g., Beck et al. 2007, Menze et al. 2007, Garrison et al. 2008, De Laet et al. 2009, Evans and Traviglia 2012). The generally rather conservative approach reflects the notion that the intrinsic variety of the archaeological record will make failure of computer-based object detection likely (see review of current debate in Cowley 2012).
While this objection may appear plausible at first glance, it often seems to be based on the misunderstanding that an automated object detection approach aims at a complete register of all archaeological sites and objects of a given region, potentially replacing visual image interpretation and/or fieldwork by archaeologists. This cannot be the goal of such an approach. Human vision is clearly superior to the current state of computer vision, and will remain so in the near future. In fact, all the above-mentioned cases of successfully applied object detection across various disciplines target only well-defined categories of objects. This is also true for the few but inspiring archaeological case studies in this field (e.g., Trier et al. 2009, Trier and Pilo 2012).

While expert knowledge and fieldwork will remain indispensable for any serious archaeological research, automated algorithms may well be used to assist with routine and time-consuming tasks. A typical example is regional archaeological survey, especially when undertaken in vast, difficult, and/or hitherto unexplored terrain. In such cases aerial or satellite images are often visually inspected prior to fieldwork with the goal to identify potential sites and typical site categories to guide fieldwork. Automated screening of large image data sets may offer valuable assistance and increase work efficiency in this case. The results would guide and enrich fieldwork rather than supersede or constrain it.

2. The Silvretta Archaeological Project

To develop methods as outlined above, one needs to start from a sample of documented sites in a region covered by remotely sensed images, as is the case in the Silvretta mountain range on the Swiss-Austrian border (Fig. 1). An ongoing research project in this area (Reitmaier ed. 2012) is currently serving as a case study to explore the potential of using object detection on remotely sensed images to assist archaeological survey.

Archaeological fieldwork in the Silvretta region was initiated in 2007 by Thomas Reitmaier, then at the University of Zurich, with a focus on the human occupation of the region during the Bronze and Iron Age. An important goal was to investigate the origins of alpine pastoralism, or *Alpwirtschaft*, an economic system of resource use highly adapted to the challenging environmental conditions in the high mountains (Reitmaier 2010, Gleirscher 2010). With the incorporation of project partners from the universities of Innsbruck, Konstanz, and Bamberg, and from such diverse disciplines as paleoecology, geography, geology, linguistics, and computer science, research in the Silvretta region has since grown into an international and interdisciplinary project. It addresses among other topics, Holocene landscape genesis and settlement history, human-environment interaction, and questions of cultural continuity in the high mountains (Reitmaier ed. 2012, Walser and Lambers 2012). Within this broad framework, the earliest indications of human impact on the landscape, and the transition from hunting to herding, defining the origins of alpine pastoralism, remain key research questions.

The permanent basis of alpine pastoralism is the settlements of farmers in the lower valleys, where livestock is kept in stables during the winter. As snow-covered areas recede in spring, livestock is driven uphill to pastures on the valley slopes and finally, during the short summer, onto fertile pastures above
the tree line, i.e. areas otherwise unsuitable for permanent occupation. In archaeological terms, alpine pastoralism is associated with built infrastructure that can still be observed in our study area today, like huts and cabins for seasonal dwelling, enclosures to protect livestock from wild animals, and, if dairy farming was involved, cheese cellars with access to water. Many of such architectural remains have been documented in the Silvretta region (Fig. 2). While the harsh environmental conditions have caused substantial damage to some sites, most are still visible on the surface since destruction by modern land use is minimal.

3. Project Goals

While being unique in their specific forms, the above-mentioned sites and buildings show a limited variability of geometric properties in their shapes and proportions. They thus constitute a typical problem of object detection, in which the task is to detect structures of similar shape. In order to develop a general approach of archaeological object detection, we decided to use the registered sites associated with alpine pastoralism, among them approximately 20 well recognizable ruins, as starting point. Corresponding to their function and purpose, all of these sites are located in open grassland, most of them in the alpine zone above the tree line (Fig. 2). We therefore defined open grassland as our region of interest.

Other categories of archaeological sites also registered during our survey, such as fireplaces and rock shelters, are not considered in this approach. Furthermore, we do not currently pursue other potential approaches we discussed prior to image acquisition to identify archaeological sites (Lambers and Reitmaier 2013). Our goal is to develop a method that assesses the probability of the presence of archaeological objects of the type described above in a given area based on their geometric properties. The expected result is a probability map of candidate sites in our study area that can be visually verified prior to fieldwork. The quality of such a map will be judged by its sensitivity to structures resembling objects of our interest and the rate of false detections. A low rate of false detections in various terrains is essential for our methodology to be useful, which means that segmentation of high contrast textured regions as described below is crucial.

It should be noted that our ultimate goal, the detection of archaeological structures in vast unknown areas, differs essentially from contrast enhancement and automated mapping of known sites and structures as undertaken in other archaeological projects (e.g., Jahjah and Ulivieri 2010).

4. Image Data

We decided to use high-resolution satellite images for image data. Although our study area is covered by various sets of aerial images, none
of these is available for the whole study area because the study area is divided between two Austrian federal states and one Swiss canton. Furthermore, in mountainous regions aerial images suffer from distortion caused by greatly varying terrain elevation. Satellite images, on the other hand, are able to cover the whole study area with consistent, up-to-date images and are less distorted due to the much greater distance between camera and terrain surface.

Another more general reason led us to choose also satellite images. We believe that satellite images with a high spatial resolution of 1m and better will prove an important data source for archaeology in the future (Parcak 2009, Giardino 2011, Lasaponara and Masini 2012). Since this kind of imagery first became available in 2000, its quality and variety has increased dramatically. Today, nearly worldwide no other high-resolution image data source is available under consistent conditions and without major legal constraints. High-resolution satellite images are therefore likely to become the preferred data source for regional archaeological research in many areas of the world where alternative data sources, such as aerial images, are not easily available.

For the Silvretta project, we chose to order GeoEye-1 images. The camera mounted on this satellite launched in 2008, captures images featuring a panchromatic channel and four colour channels (RGB + NIR) with a spatial resolution of 0.41m (pan) and 1.64m (VNIR), respectively. We ordered the bundle product comprising the panchromatic channel and pansharpened colour channels. Due to legal regulations, after pansharpening all channels were downsampled to a spatial resolution of 0.5m pixel size. While lower than that of most aerial images, this spatial resolution still allows the detection of structures pertaining to our target objects. The characteristic size of the archaeological ruins we are looking for varies roughly between 10 and 100 pixels. Their walls are generally not wider than two pixels. Reliable detection of structures of a few pixels in width is very limited. In some cases it might not be achievable at all. However, in our case the second dimension of walls, whose length is usually above 10 pixels, makes this task still possible.

For improved georeferencing and orthorectification of the GeoEye-1 images, we provided ground control points measured by graduate students of geomatics at ETH Zurich during their summer field course in June 2011. Furthermore, we ordered images with a cloud cover of less than 5%, which is important in the Alps where weather conditions are volatile. After a rainy summer, four scenes of our study area where finally captured on September 6, 2011. Following image processing by GeoEye, we received the image portions that cover our study area of approx. 540 sq. km in early November 2011. By that time, we had already started to develop first elements of our method on Swisstopo orthoimages with the same resolution of 0.5 m to expedite the start. These elements were applicable to the satellite images, too. All further research was conducted on the GeoEye images.

5. Digital Image Analysis: General Methodology

Our method aims at capturing geometrical properties of the objects of our interest mentioned above, namely ruins of livestock enclosures and alpine huts. Such structures can be modeled by linear features that meet at approximately right angles.

Our general methodology is divided into several stages that extract image features of growing size and complexity. In the first stage, local features such as black and white blobs stemming from the background are extracted from the image. This stage can be implemented by means of white or black top-hat transforms (Soille 2003) or their combinations. Colour information can also be incorporated into these transformations, as suggested by Hanbury.
Towards Detection of Archaeological Objects in High-Resolution Remotely Sensed Images
Karsten Lambers and Igor Zingman

While we currently do not use colour information, this could be beneficial in cases where colour contrast is higher at the image features we are interested in.

Chains of blob features with possible gaps in between may form linear features. During the second stage, we group extracted blobs into larger, approximately linear features that may correspond to the walls of ruins of huts and enclosures. Near right angle intersections of linear features that may be defined by corner points, are then searched in the next stage. Although many approaches for line and corner detection exist, they are not robust enough to be widely applied to different types of images with different scenery, textures, irrelevant structures, noise, illumination conditions, and contrasts. We are currently investigating these topics, aiming at the development of new line/ corner detection algorithms adjusted to our specific task and the type of imagery we use.

Evidence of structures of interest can then be inferred from extracted linear features, corners, and contextual keys, if available, for instance surrounding texture. This is the final stage of our methodology to be developed in the future. The output of the last stage is a probability map indicating the presence of objects of interest. Such a map will have zero values at most regions, and a continuous range of probability values at other regions. This map can be further thresholded at the level corresponding to an acceptable rate of false detections.

Our database that includes about 20 objects of our interest is rather limited when taking into account the variations in the appearance of the objects. It thus does not allow us to employ machine-learning techniques. It nevertheless allows us to develop approaches based on geometrical properties and other known key features. It also allows for a coarse estimation of the achieved sensitivity. Since we acquired a large amount of satellite images, we can reliably test the corresponding rate of false detections.

For our general approach to succeed, it is essential to discriminate between smooth and high contrast texture regions in our images. For this basic prerequisite, we developed a new method described below.

6. Texture Detection

While the concept of texture has no precise definition, it is generally associated with repeated changes in image grey level. In remotely sensed images such as the ones we use in the Silvretta region, high contrast textures are, for example, forests (Fig. 3) or urban areas. Large amounts of local features are usually extracted in texture areas in the first stage of the described methodology. This is because the local operator does not distinguish isolated features like blobs or lines from features that belong to a different context, a texture. Grouping at the following stages does not suppress these features since they are easily grouped with other surrounding texture features resulting in unexpected false detections.

To overcome this problem in areas such as forests, urban areas, or rocky mountains, we developed a new texture detector that
filters out high contrast textured regions irrespectively of texture type. Since the objects of our interest are located in open grassland, i.e., outside the above-mentioned areas, this procedure reduces the number of false detections without affecting sensitivity to true examples. Filtering out textured areas also greatly reduces computational burden since the developed texture detection algorithm employs a sequence of morphological operators that are much faster than algorithms for the detection of geometrical structures of interest.

7. Morphological Texture Contrast (MTC) Transformation

The approach we introduced recently (Zingman et al. 2012) aims at detecting high contrast texture regions of different types. It is based on mathematical morphology, which has proven to be very efficient in the processing of remotely sensed images (Soille and Pesaresi 2002). We call our approach Morphological Texture Contrast (MTC) transformation.

In comparison to many other texture detection approaches developed to discriminate different types of texture, the MTC transformation is insensitive to texture properties except of texture contrast. It is intended to discriminate smooth regions corresponding to our regions of interest from regions with high contrast texture, such as forests, urban or rocky areas in remotely sensed images. An essential property of our detector is the ability to provide a low response at isolated or individual features, even if they are of a high contrast, a result that is currently not achieved by other techniques.

A commonly used technique to detect texture is to measure the local standard deviation of grey level within an analysis window moving over the whole image (Gonzalez and Woods 2001). We refer to this technique as a variance-based technique. It produces a response proportional to the texture contrast, thus allowing texture regions to be detected. However, it also produces a high response at isolated features (Fig. 4a), which in our case
is unacceptable. Another disadvantage of the variance-based measure is that it blurs the borders of texture regions, which may result in inaccurate texture localization (Fig. 4c). Modifications of the variance-based technique are frequently used to generate a contrast feature within a set of other descriptors to discriminate textures (see, for example, Ojala et al. 2002).

The MTC transformation was developed as an alternative approach based on morphological alternating filters, namely morphological closing followed by opening and opening followed by closing (Serra 1988, Soille 2003). These operators are usually used for the suppression of noise in images. We use these operators to build a transformation proportional to texture contrast. The MTC transformation is obtained by taking non-negative values of the difference between morphological closing followed by opening and opening followed by closing. Negative values of the difference are substituted by zero values at the output of the MTC. This transformation results in high response at textured areas and low response at isolated features and in smooth regions.

While in our project the MTC transformation is applied to 2D images, the underlying idea is more easily explained by using an artificial 1D signal as shown in figure 5. The MTC transformation measures the difference between the upper and lower envelopes of texture providing a response proportional to texture strength. These envelopes, obtained using alternating morphological filters, coincide at isolated features, resulting in a suppressed response at such features. This capability of the MTC transformation is not available in other approaches.

Figures 6c and d illustrate the MTC transformation applied to 2D aerial and satellite images of 0.5m resolution shown in figures 6a and b. High response, corresponding to bright grey values, is generated in texture areas, while low response is produced in smooth areas and at isolated features. Low response at isolated features, such as individual trees, can be more clearly seen in the zoomed area shown in figure 4b that was cropped from the left corner of figure 6c. Figure 4d emphasizes how accurately texture regions are aligned with the regions of high values of the MTC descriptor.

The distribution of grey levels in the transformed images is highly bi-modal, with one mode corresponding to texture regions (high grey tone values) and the other to smooth regions (lowgrey tone values). These two modes can easily be separated by finding an appropriate threshold. This provides us with a segmentation that defines two disjoint masks for texture and smooth areas. We used the Otsu thresholding method (Otsu 1979) to find an appropriate threshold automatically. This approach corresponds to an unsupervised classification scheme since input data does not need to be labeled. The segmentation results in figs. 6e and f are superimposed on the initial images to emphasize the alignment of the results with the original data. As can be seen, the segmentation is quite accurate at the borders of texture regions.
Figure 6. (a) Pan-chromatic image of 4000x3500 pixel size and 0.5m pixel resolution captured by the GeoEye-1 satellite (© GeoEye 2011, distributed by e-GEOS). (b) Aerial SWISSTOPO image of 6100x5000 pixel size and 0.5m pixel resolution. Scenery in both images includes high contrast textured regions (urban and forest areas) and comparably smooth field areas. (c) and (d) The MTC descriptor applied to both input images. (e) and (f) The segmentation result superimposed on the original images as obtained by automatic thresholding of the MTC descriptor. Brownish areas correspond to high contrast textured regions.
Towards Detection of Archaeological Objects in High-Resolution Remotely Sensed Images
Karsten Lambers and Igor Zingman

In cases of highly varying sizes of texture and smooth areas in a given image, one of the two modes of the histogram (of the image transformed by the MTC operator) may be too small to reliably find an appropriate threshold. In such cases, the user can choose to use a supervised approach in which they are asked to manually delineate several representative regions of texture and smooth areas. Then a simple supervised classification technique in one dimension may be employed, e.g., the nearest distance (nearest mean) classification (Duda et al. 2001). Since manual delineation does not need to be accurate, it can be performed easily and quickly by a human operator.

So far, our tests showed the MTC transformation to be relatively fast. An image of 6100x5000 pixels is processed in about 21 seconds by our code written in Matlab installed on a standard PC equipped with an Intel Core2 Quad 2.83 GHz processor. This processing time corresponds to a square analysis window of 40x40 pixels. Our technique does not require parameter tuning except for a single scale parameter that defines the size of the analysis window. This parameter should just be roughly adjusted to the characteristic size of texture. For remotely sensed images it is related to their spatial resolution and the distances between objects on the ground. The technique is robust to illumination changes within the image and also works well with images from different sources. Though our technique was developed to analyse remotely sensed images, its application is not limited to this type of data.

8. Summary and Outlook

We have identified about 20 sites in the Silvretta region with ruins associated with alpine pastoralism. These sites serve as representative examples for the development of automated methods to detect similar sites in high-resolution satellite images. Using a texture segmentation technique that we recently developed, we segment the images into regions of interest and other regions based on texture contrast. This step is a prerequisite for object detection for which further steps have yet to be developed. By filtering out textured regions where no archaeological objects are to be expected, our approach will greatly reduce false detections.

In general terms, our approach follows successful examples of object detection in other fields by targeting a well-documented category of objects. While our sample is small, it can be used to develop methods that will later be tested and refined in other contexts. Our approach at image segmentation translates archaeological categories — regions where the archaeological objects we are interested in occur, and other regions where this is not the case — into categories of image properties, such as smooth image segments, and image segments with high texture contrast, respectively. While this approach may not always be as straightforward, in our view it is the direction to take in archaeological image analysis.

We are currently working on the method of actual object detection, the general outline of which is explained in section 5. While the outcome cannot fully be predicted at this time, we think it is important to proceed in this direction. Fieldwork in the Silvretta Alps will continue in the coming years. We will thus have sufficient opportunities to check the results of image analysis in the field. Once a functioning tool is available, it will be possible to test it in other regions with similar conditions through our network of cooperation partners working in alpine archaeology.

Acknowledgements

This project was jointly funded by the EU programme “Interreg IV Alpenrhein – Bodensee – Hochrhein” as well as the Zukunftskolleg and the Graduate School “Explorative Analysis and Visualization of Large Information Spaces”, both at the University of Konstanz. We thank
Thomas Reitmaier, Christoph Walser, and Dietmar Saupe for their helpful support, two anonymous reviewers for their insightful comments, and Anna Dowden-Williams for her skillful proofreading of our manuscript.

References


Towards Detection of Archaeological Objects in High-Resolution Remotely Sensed Images
Karsten Lambers and Igor Zingman

Cambridge Scholars Publishing.


1. Introduction

The benefits of aerial photography in archaeology are manifold, but in most cases it is difficult for single researchers or small-scale enterprises to run a project with remote sensing techniques, due to logistical problems, high cost and the complexity of the necessary equipment. Hiring microlight aircraft can be quite cost-intensive besides the difficulty of taking zenithal pictures during low-altitude flights, while balloons are consuming noble gases which are subjected to severe safety regulations for storage and transport. Kites depend too much on wind to be considered a reliable instrument in any circumstances and radio-controlled helicopters or aircraft, with internal combustion engine, can be dangerous (explosion hazard in case of crash), noisy and require routine maintenance.

Combined use of open source hardware and software can be one possible solution to the problems described above. The case study of Monte S. Martino ai Campi di Riva del Garda (Italy) allowed us to complete an aerial archaeology project in difficult logistical throughout conditions, achieving satisfactory results and maintaining a low cost approach.

For a better understanding, before reporting our experiences in S. Martino, we have to describe briefly the two main components of our remote sensing platform (ArcheOS and UAVP) and the historical and archaeological characteristics of the site.

2. ArcheOS

ArcheOS (Archeological Operating System – www.archeos.eu) is a GNU/Linux distribution developed for archaeological aims...
and released by Arc-Team under GPL (General Public License). The system contains a selection of different applications which try to cover a wide range of archaeological needs.

Created in 2004, the ArcheOS-project was first presented to the scientific community during the annual meeting “Archäologie und Computer” of the city of Vienna (Bezzi et al. 2005). From the beginning, the main purpose was to introduce and to spread the use of Free/Libre and Open Source Software and to apply the ideology of the “Free Software” movement to archaeology itself. A central objective of the Free Software Foundation is the unimpeded circulation of data, knowledge and ideas. This is, in our opinion, also an essential condition for the better development of archaeology as a discipline. That is the reason why Arc-Team is suggesting a model of data-sharing inspired by the Open Source and Free Software movement.

The first release of ArcheOS (codenamed Akhenaton) was based on the GNU/Linux distribution PcLinuxOS 0.91 (beta). The selection of software and applications (all protected by GPL or OSI compliant licenses) was the result of many years of working experience and tests conducted by the Arc-Team company. The restrictions imposed by the General Public Licence give the system all the benefits of free software: ArcheOS is modifiable, redistributable, supported by the net community and moreover it is free (gratis). Technically, ArcheOS is a Linux-Live DVD which can be used “on the fly” by inserting the disk before booting the PC. In the live-mode the whole operating system will run using the RAM memory and no permanent changes will be applied to the computer. At the end of the session, the native operating system can be restarted without any modification. The live-mode is mainly suitable for testing the distribution, which can always be installed permanently after partitioning the hard-disk.

In 2006, the ArcheOS project experienced a lightning development thanks to the contribution of the Aramus Excavations and Field School (Kuntner and Heinsch 2008) of the University of Innsbruck (Institut für Alte Geschichte und Altorientalistik), which shifted its hardware to the free operating system. The field-work experiences in Armenia were used by Arc-Team to further develop ArcheOS (Bezzi et al. 2006). The feedback of Aramus Field School has been very important for building the new version, giving us the opportunity to consider the specific needs of an archaeological mission abroad. Moreover the possibility to share knowledge and software with students increased exponentially the community. From this collaboration started the Digital Archaeological Documentation Project (http://vai.uibk.ac.at/dadp/), a wiki system to publish on-line tutorials connected to the use of FLOSS and ArcheOS.

Thanks to the Aramus experience, a new version of ArcheOS was ready in 2007. The main features of the second release (codenamed Sargon) were the migration to the Kubuntu GNU/Linux v. 7.10, which ensured a better hardware integration, and a more accurate selection of programs, which derived directly from the hard test conducted in 2006. ArcheOS’ historical timeline led in 2009 to the third release (codenamed Xenophon). The base GNU/Linux system was Kubuntu v. 9.04 which integrated the user interface KDE 4. The first three versions of ArcheOS were not upgradable: the only way to update most of the applications was to install the new release deleting the old one. In 2010 the computer scientist Fabrizio Furnari joined the developers team. He started to reorganize the structure of the project, paying attention to the stability of the operating system and to an on-line service to keep up to date with the different programs (deb-repository). As a result, in 2011 Arc-Team was able to share the current release (codenamed Caesar) based on Debian Squeeze v. 6.0.

ArcheOS tries to satisfy all the needs of an archaeological project, covering every single step of the operating workflow, from
data collection and storage to elaboration, publication and sharing. The software selection is organized by typology in different menus (Fig. 1). For instance the submenu GIS includes, among other applications, the powerful software GRASS, which integrates several high level analysis tools, QuantumGIS, an intuitive raster-oriented platform, and OpenJUMP, which is valued for its vector drawing potentialities (very close to CAD). The Graphics submenu offers different choices for raster or vector datasets: raster images can be edited using GIMP, while a vector drawing can be realized within Inkscape. Moreover Blender can be used for virtual reality projects (3D), while the application Stippler allows the user to turn a greyscale picture into “stippled” images (following the most common conventions of the archaeological drawings techniques). CAD software are represented by LibreCAD, for architectural 2D drawings, and FreeCAD, for parametric 3D modelling. The database submenu is filled with a wide range of applications, from the simple OpenOffice Base to the complete object-relational DBMS PostgreSQL (implemented with the geographical support PostGIS). In addition the new release of ArcheOS has been enriched with SQLite and SpatiaLite. Moreover, other peculiar professional fields are supported by specific software. For instance MeshLab represents the “killer application” for mesh editing, while Python Photogrammetry Toolbox (Bundler+CMVS/PMVS2) is an optimal solution for acquiring 3D point clouds with Structure form Motion and Image-Based Modelling techniques. Scientific 3D models can be imported inside Paraview, a multi-platform data analysis and visualization application, as well as in VirtualTerrain Project, a 3D viewer for geographical datasets. In most cases statistical analysis are delegated to the powerful software R. Obviously ArcheOS comes also with an office suite, many internet tools and multimedia applications.

Being an open source project, ArcheOS is continuously under development, thanks to the support of a community of users/developers which shares knowledge, data and experiences to improve the quality and the usability of the operating system.

3. UAVP and Other Flying Drones

UAVP is an acronym for Universal Aerial Video Platform and refers to a project whose purpose was to build “a flying Open Source QuadCopter at a reasonable price”. Like many other open source projects, the UAVP has grown rapidly. In fact the original device, born from the initiative of Wolfgang Mahringer, was followed in recent years with a community-driven design, which led to the new prototype: the NG-UAVP (Next Generation UAVP). From a technical point of view an UAVP model can be counted in the category of UAVs (Unmanned Aerial Vehicles), being a flying radio-controlled drone. Its main strengths, in the field of aerial archaeology, are connected with its stability and its capability to hover in place and are derived from its “quadcopter” structure. Also its electric power supply can be considered a benefit for safety during working operations, even if offset by disadvantages in terms of flight range.

In 2008 Arc-Team built its first UAVP prototype (Bezzi et al. 2008) based on the original Wolfgang Mahringer project (Fig. 2). Considering the complexity of this kind of flying machine, normal archaeological vocational training doesn’t offer the specific
skills needed to assemble and install the different parts of the hardware and software. For that reason collaborating with an expert in modelling (or an electronic engineer) is strongly recommended to successfully build and develop an UAVP. There are no particular difficulties in assembling the mechanical parts (frame), while the electronic components are effectively more sophisticated and require a minimum competence in soldering circuits and a basic knowledge about concepts like motherboards and processors. However the basic elements are few (many are optional) and it is possible to buy some preassembled SMDs (Surface Mounting Devices), like the motherboard (available already populated). The final prototype is composed of a carbon fibre frame and a support for a digital camera, equipped with an additional servomotor for its orientation. The electronic component is restricted to the essential elements: motherboard, gyroscopes, processor, linear acceleration sensor and electronic compass.

To experiment different solutions, in 2010 Arc-Team has built a KKmultiCopter drone (Fig. 3). This project is a simple set of hardware components already populated, developed by the public domain KapteinKuk model. The electronic elements are sell in a all-in-one solution integrated directly into the motherboard. The project is supported by a wide net community, particularly active in hardware/software development, and by several on-line resellers. After a long training (it is important not to underestimate the necessary time to learn how to pilot a quadrotor) the KKmultiCopter drone can impress for the stability of flight and the manoeuvrability.

4. The Site of Monte S. Martino ai Campi di Riva del Garda (Italy)

The archaeological site of Monte S. Martino is located in the area to the north of Lake Garda, about 800 metres above sea level. Its strategical position was connected to a network of several routes linking the Brescia territory (to the west) and the Alpine area (to the north). Only in 1880 did the site start to arouse the interest of the scientific community, but during the second half of the 20th century its territory, which was once used as pasture, was occupied by dense vegetation and did not preserve traces of remains on the surface. In 1969 a group of enthusiasts noticed the presence of numerous fragments of Roman throughout tiles and decided to start an excavation. Six years later the site came under the protection of the Provincia Autonoma di Trento: this was the beginning of systematic research in Monte S. Martino.

Until now the oldest archaeological finds brought to light so far at the site are represented by a blade of stone axe and a flint arrowhead, both related with prehistory and both discovered by chance. The axe blade probably
dates back to Copper Age (third millennium BC) and it is made of jade, a valuable raw material that comes from the Western Alps (Piemonte or Liguria). The flint arrowhead can be dated between the early stages of the Late Neolithic and the Early Bronze Age and it may have been lost during hunting activity (Mottes 2007, 163-164).

Almost all the pre-Roman artefacts from Monte S. Martino belong to the second Iron Age (VI - I centuries BC). The characteristics of these findings show an interaction between the two main cultural environments prevailing at the time, the Group of Fritzens-Sanzeno and the Group of Valcamonica (Marzatico 2007, 172-173). Probably the people who frequented the area during this period practised rituals similar or somehow connected to the so-called Brandopferplätze. In fact the excavations conducted in the highest area of the site revealed an archaeological record with layers containing a considerable percentage of carbon and significant traces of combustion together with fragments of pottery cups, probably crushed as cultural offerings.

Other traces of the pre-Roman period were found all over the site and recently emerged in the south-eastern sector, below a settlement dated between the IV and the VI Century BC. New structures were discovered in this area, exposing some parts of a square building, bordered by stone walls.

It seems that the identity of pre-Roman population was very strong and it partially survived after the arrival of the Romans. Since the end of the first century BC, a shrine was built in the same place of the pre-Roman rituals. The structure has been used until the third century AD and presents a rectangular plant (Fig. 4), which surrounds the highest part of the area. It is divided in several rooms, adapted to the terrain morphology and probably connected with the practical needs of the sanctuary. They were initially built only in the south-west area, but later other structures were added along the east side, giving the complex the current appearance. The building entrances were probably located on the West, South and East side, while the North side overlooks a precipice. The sacred function of the site is shown by the artefacts found in some collapsed rooms or in the proximity of the sanctuary. In this area were discovered fragments of three altars, two of which bear inscriptions which testify to the strong relationship between natives and Romans. In fact one of this epigraph is epichoric, being written in a local language, but employing the Latin alphabet. The second inscription is in Latin, but shows indigenous names (Valdo 2007, 343-346).

It is assumed that many different deities were worshipped in the shrine, and that the ancient indigenous beliefs were merged within the Roman Pantheon (Bassi 2003). The sanctuary, with its location in the territory of the Upper Garda Lake had to have a significant religious importance in the local area.

In the fourth and fifth centuries, numerous dwellings were built in the South-eastern area of the sanctuary, using artificial terraces. These structures are mainly built of stones and mortar and are partially excavated.
in the mountain rock. In the same area has been brought to light a building that stands out for its size. This characteristic, together with its position, is compatible with an hypothetical public function. Moreover, the village seems to be surrounded by fortified walls. From this point of view, the defensive configuration of the settlement, as well as its strategic position, allows to advance the hypothesis (to be verified) that it could have been a military garrison with logistical purposes. Anyway, all the investigated buildings were abandoned in the sixth or in the beginning of the seventh century, destroyed by fire. Some of the structures, however, were probably used also in later periods, like the one in which was found a coin of the Byzantine emperor Heraclius (610-641), located a few dozen of meters in North-west direction from the original nucleus. Perhaps already in the sixth century a small church abutted on the west wall of the public building, which became the front of the religious structure (Bellosi et al. 2010). During the twelfth or thirteenth century, the church was renovated: the apse was heavily reinforced, while the front wall was advanced westward; also the floors were redone and paved with stone slabs which raised the walking level. In this time the churchyard continued to be bounded by the South wall of the ancient public building (abandoned). In the proximity of the church entrance, a small structure and a tank probably served a building connected with craftwork productions. The area outside the church has been used as a burial place. The building was interdicted in 1612, as recalled in a document dated 1750. This paper decreed the destruction of the religious structure because of its poor condition. It is not known if the order has been executed, or if the church was abandoned until it was completely buried, but the final result was the abandonment of the site, which initially became a pasture and then turned back to wood, until was rediscover with the excavation of 1969.

5. The Remote Sensing Experience in Monte S. Martino

During the winter 2012 Arc-Team was hired to take aerial pictures of the buildings of Monte San Martino. Thanks to the difficult logistical conditions of the site (mountain area, arboreal vegetation, wind) it was a perfect situation to verify the effectiveness of the UAVP model. The strategy applied on the site included different low-altitude flights in four areas free from vegetation cover (Fig. 5), both in First Person View (FPV) and in Direct Visual Flight (DVF). All the project was supported by the advice and the practical help of two expert aeromodelers (Walter Gilli and Walter Morelli).

For the first flight, the drone was equipped with two different digital cameras: a Nikon CoolPix S210 and a GoPro HD Hero. The first device is a compact-camera with a 8.0 MP sensor, an angle of view of 64° and an aperture range of f3.1-5.9. The shoot signal is manually sent by a remote control. The second instrument is a sport-camera designed for extreme conditions. It has a 5.0 MP sensor, an angle of view of 127° and an aperture size fix to f2.8. The internal software can automatically take three pictures each second. Both the cameras have some benefits and some disadvantages, but finally the best results were obtained using the GoPro (Fig. 6). Despite a general low quality of the images, all the pictures taken with this camera were in focus (thanks to the fix aperture f2.8)
and included the whole area of interest (thanks to the wide angle). The high lens distortion did not represent a problem, being removed during the post-processing phase. On the contrary, in windy condition, the Nikon camera was hard to manage: the manual shooting was a cause of problems in piloting the drone and the slow shutter speed, due to smaller diaphragm’s aperture size, produced many images out of focus. For this reason, the other flights where supported only by the GoPro camera.

The results were partially processed directly in the field using a rugged PC and a total station. The image elaboration was done with the software Python Photogrammetry Toolbox. At the end of the process all the pictures were undistorted (lens correction) and 3D point clouds of each area were created (Fig. 7). The undistorted images were rectified and georeferenced using the “Metodo Aramus”, a combination of different software (GRASS, GIMP and E-FOTO) which allows the collection of good quality photomosaics, while 3D point clouds were transformed in mesh within the application MeshLab.

After the Monte S. Martino experience, new tests are planned to further develop the whole system. A possible implementation can be the integration of the new GoPro HD Hero2, which should solve some problems experimented with the old model. The 11.0 MP sensor, the variable angle of view (170° and 127°), the faster shutter speed and the automatic shooting are features that makes this camera a perfect device for aerial photography data acquisition through RC drones. Moreover a new frontier for UAV models is the integration of GPS navigation tools which allows to program autopilot flight and to maintain a desired position. The first tests conducted with the new configuration of the remote sensing platform are giving promising results.

References


The Visualization of the Archaeological Information through Web Servers: from Data Records on the Ground to Web Publication by Means of Web Map Services (WMS)

Julio Zancajo and Teresa Mostaza
USAL, E.P.S., Spain

Mercedes Farjas
UPM, E.T.S.I. Topografia, Spain

Abstract:
The purpose of this study is to show the processes of data record and their structuring in Spatial Database, for the integrated management of the information associated with an Archaeological Site, and their visualization using web servers, by means of developments based on free software in order to minimize costs. Regarding spatial information, new measurement systems have been created over the last few years. They provide new approaches to the surveying works and complement the spatial documentation processes of the Sites. The developed system will allow the integration of the new information generated by these systems, as well as the visualization of new products, whether spatial or not. Moreover, it will allow the integration of georeferenced information about the Archaeological Site together with other cartographic documents, such as Orthophotographs, so as to get a complete spatial visualization. Finally, the objective is to use software map servers Open Source (MapServer, GeoServer), and by XML programming, to generate a web visualization environment, friendly and easy, based on OGC standards.

Keywords:
Heritage, SDI, OGC Standards, Open Source Technology

1. Introduction

We started working on archaeological sites four or five years ago. Initially, we did the tachymetry of these sites, with two objectives:

• Obtaining the geometry of some Archaeological elements.

• Georeferencing the archaeological site for the integration with other cartographic documents, because today the government demands this georeferencing.

After that, we started to structure the archaeological site information to manage with Geographical Information Systems (GIS).

In order to structure the information of an archaeological site properly (Aburizaiza 2009; Campagna 2006; Steinberg 2006), it needs to be taken into account that there are two types of information:

• Firstly, purely archaeological information, related to the information provided by the site-structures, pottery, coins.

• Secondly, spatial information, this including the spatial absolute position of the Archaeological Site, the relative position of the elements discovered inside, as well as the dimensions and geometries of the Archaeological Site itself and the objects discovered.

On the other hand, today, there are instruments that they offer new possibilities in the documentation of archaeological sites. The same instrument may be to offer results metrics (cartographic) or not metric documents.
The Visualization of the Archaeological Information through Web Servers
Julio Zancajo, Teresa Mostaza and Mercedes Farjas

depending of the processes applied, apart the documentation provide for the traditional archaeological jobs.

This improves the processes of documentation of the archaeological sites, because of with the same effort we can obtain different products, everything very important to understand and to improve the analysis process (Zao 2011). For example, the UAVs can provide through photographs a strictly metric document, the orthophotography, with a big resolution (Fig. 1), or a non-metric document, in this case, an aerial panoramic, which, however, may be an interesting visual document to know the archaeological site (Fig. 2). This only depends of the processes applied.

2. Aims of Metrical Instruments

The metrical instruments, together to application metric methodologies, provide two types of information:

- The metric information of the archaeological objects. This metric information is very important for different process, as for example, for their future reconstruction or, more frequently, to study them without use the original. This second application allows to researchers to work with fragile objects, which could spoil if we work with theirs. The best option for these cases is to work with a virtual model, which can obtain trough of the metric information.

- The georeferencing information of the Archaeological Site. This allows integrating the cartographic information of the Archaeological Site next to other cartographic documents (for example, the Ortophotography) to its management by GIS and its visualization on Web Map Services (WMS), in application the INSPIRE directive, using the OGC standards (OpenGIS 2011).

Today, this second application is very important, for example, to display the results of the research with the spatial dimension integrated as an aspect of interest to learn about spatial interactions in the territory, and in this paper we study the problem of the integration and the management of the archaeological information.

3. The Management of the Archaeological Information: Problems and Adopted Solutions

The spatial information does not have any problems to manage with spatial technologies (GIS or SDI), because these system are the suitable for their management.
The problem appears with the archaeological information which source is the traditional jobs on the archaeological sites, because this information can be non-spatial and its management could be difficult into spatial systems, because this information can not be displayed on maps and the GIS software with typical tools does not provide a management easy for this type of information.

However, it is necessary the integration of this type of information into the spatial database, because this information is necessary for the archaeologists and it is part of the Database of Archaeological Site. For these reasons, the management of archaeological information is incomplete in many systems:

- Archaeologists use GIS to display the geometries, sometimes with a few attributes for specific analysis.
- The rest of archaeological site information is managing with other systems, for example, for multimedia files, pdf files, etc.

This situation is not the best because it does not allow working with all information into a single system and it can produce problems in the Database. The objective is to manage the information with a single system, regardless of the source of information.

3.1 The Management of Heritage of Ávila (Spain)

The research to find a solution to the problem raised started three years ago. The “Gran Duque de Alba” Institution in Ávila (Spain) manages the Cultural Heritage in the province and it needed a system to manage the information associated to the Cultural Heritage in Ávila.

We started to develop a system to manage the information of the Heritage, under the following conditions:

- The system would work on Microsoft Windows and it would be based on the GIS System of Intergraph, Geomedia.
- The development would be a specific interface with custom tools for workstation, allowing the information management to non-expert GIS users.

We started the study of the information of the Heritage and we designed a structuring of the Spatial Database and the interface for the system (Fig. 3).

We also decided the cartographic documents which the system would use as georeferencing the Cultural Heritage Information. For example, as general display we used the province map at scale 1/200,000, due to the high information of this map (Fig. 4).
However, this system is not a complete solution, because does not allow to manage of information without spatial definition. Moreover, the system allows the access to information which could not be into the Spatial Database. To solve this problem we proposed a field as link web which allows accessing this type of information or the information stored on web. The Database is very simple and it also allows seeing more information about the entity, for example, 3D models (Fig.5).

The second step was the development a system to visualization of the Heritage on Web (Fig.6). The web system is based on Geomedia WebMap (software of Intergraph) and the conditions are similar to workstation system:

- Direct access to main information of elements of Heritage.
- To allow the access to another information (Fig.7).

At the moment, we only can access to the information of the Avila Township. The Web direction to access to the Avila Heritage information and to explore the different options is http://sig.diputacionavila.es/inicio.

4. “Archaeologica”: a Proposal of Integral Solution

The systems described in the previous section are based on commercial software and do not solve the problems associated with the management of the Heritage information. To solve these problems, we started to work about a new system under the following conditions:

- The system should be based in Open Source software. This condition is essential not to increase the costs (licenses and later maintenance) (Garbin, 2010; Quian Yi, 2008; Stitfer, 2009).

- The aim is to use a single system to manage and display the Heritage information, included on Web, to satisfy the INSPIRE directive.

- The system should use the OGC standards.

The aim is to give an integral solution for Heritage information, from data collection to Web publication, using the OGC standards. This will allow the connection by Web, using other GIS systems or web browsers.
The Figure 8 shows an image of the workstation version, which development is almost complete. The main options are working, and we have tried to implement the work concept.

To test the system we are working with the Heritage information of Ávila (Spain). The archaeological site is the Avila city, where there are archaeological elements scattered throughout the city. Really this is not a typical archaeological site, but this is good for the test. The Figure 9 shows the distribution of the elements over the orthophoto of Avila and the access to information. We can see that the window has three panels:

- The first panel shows the attributes.
- The second panel contains the information about the spatial definition (point, line or area) and the path of file.
- And the third panel allows to access at the other documents associated to the entity (pdf files, multimedia files, etc.).

Moreover, it is possible the customization of the system to solve the specific situations to update easily the Heritage information, for example, by dialog boxes with predefined options. This avoids many problems associated with the update processes of the database.

Finally, the system provides an easy way to publish the information on WEB, based on OGC standards. Clicking on the button (Fig. 10), the system allows generating the XML file to use with MapServer to share the information on WEB, by a web browser (GetMap Request) or by another GIS software, and display this information next to their own information.

5. Conclusions

We have shown that using Open Source software is possible to manage the Heritage information, developing a specific interface to manage and display this information, although it is necessary customize the interface for specific processes, for example, updating information. However, the different systems will share the 95 % of the code, probably. This makes that really the application is one, but it is possible to perform changes to adapt it to specific needs.

Acknowledgements

This work has received financial support from the Spain Government (MICIN,
HAR2009-14248-C03-01). All this support is gratefully acknowledged.

References


Theoretical Approaches and Context of Archaeological Computing
Crafting Archaeological Methodologies: Suggesting Situational Method Engineering for the Humanities and Social Sciences

César Gonzalez-Perez
Institute of Heritage Sciences (Incipit), Spanish National Research Council (CSIC), Spain

Charlotte Hug
Université Paris 1 Panthéon-Sorbonne, France

Abstract:
Archaeological projects vary enormously in size, complexity, object of study, timescale and other characteristics. Finding the best methodology for any given project is often difficult, and an inadequate choice can ruin many months’ worth of fieldwork, bias data interpretation, and slow down or impede cross-project comparison of results. Ideally, a methodology should be adjusted to the needs of the project, based on well-proven techniques and knowledge, clearly expressed for the sake of enactment and communication, and enhanced over time. Situational method engineering has been successfully used in some disciplines to achieve these goals. This approach does not conceive a methodology as a monolithic black box, but as an assembly of pre-existing components that are selected from a repository and composed together. Each component encapsulates a proven, reusable and self-contained “atom” of knowledge. This paper offers an introduction to situational method engineering from the perspective of archaeology, showing how archaeological methodologies can be constructed following its principles, plus an analysis that shows why it is expected to be useful in this context.

Keywords:
Methodology, Process, Situational Method Engineering

1. Introduction

Most efforts in the confluence of information technologies and archaeology have been applied to the organisation, systematisation and overall betterment of data. As testimonial proof, of the seven major thematic blocks in the 40th Computer Applications and Quantitative Methods in Archaeology conference (Southampton, UK, 2012), only one (“Theoretical Approaches & Context of Archaeological Computing”) is not extremely data-centric.

However, information technologies have been successful outside the realm of archaeology, in organising and improving processes as well as data. Process-centric approaches have long been crucial in areas such as human-computer interaction, business process engineering and reengineering, quality assurance in software systems and requirements engineering. In addition, process-centric approaches are key to the work in the field of methodologies, which is the focus of this paper.

At the same time, recent debates in the field of archaeology suggest that the formal treatment of processes, often neglected, is becoming more and more present in the research arena. The Digital Research Infrastructure for the Arts and Humanities (DARIAH, http://www.dariah.eu), an ambitious EU-funded initiative, establishes goals such as “To support the cross-fertilisation of ideas, methods and expertise”, “To facilitate the development and sharing of expertise, tools, and ICT methods for the creation, curation, preservation, access and dissemination, and use of humanities and cultural heritage digital information”, or “To increase ehumanities capacity by developing, promoting and supporting digital scholarship and the application of advanced ICT methods” (DARIAH Consortium 2009). In consonance
with this, some works carried out at the Digital Curation Unit of the Athena Research Centre (Benardou et al. 2010; Constantopoulos et al. 2009) attempt to formalise the scholarly processes that are part of arts and humanities specialists. Others efforts, for example, focus on “workflow protocols” that archaeologists employ in order to obtain requirements for tools (Fabricatore and Cantone 2008), model the process to read ancient texts (Terras 2005), or present different types of process models that can be used to represent research processes in humanities (Hug et al. 2011). Finally others, such as (Djindjian 2012), claim a bigger role for process (as opposed to data) in archaeological computing applications.

Going beyond the apparent signals that process is in need of some attention, the fact is that there is a significant body of research on methodologies within the field of information technologies that, hypothetically, can be successfully transferred and applied to archaeology, and to the humanities and social sciences in general. In particular, we argue that the methodological approach named situational method engineering (Kumar and Welke 1992; Brinkkemper 1996; Rolland and Prakash 1996) is an excellent candidate to fulfil the demands that the community is making about process and methodologies.

Section 2 below explains why a formal treatment of process and methodologies is necessary in archaeology; section 3 gives some background details on method engineering itself, plus a specific proposal on how to apply it to archaeology and the humanities and social sciences in general; and section 4 presents a sample scenario that illustrates a fictitious application of situational method engineering in an archaeological context.

2. Motivation

A common definition of “methodology” is the “specification of the process to follow together with the work products to be used and generated, plus the consideration of the people and tools involved, during an information-based domain development effort” (ISO/IEC 2007, section 3.2). This means that methodologies are concerned about the process (i.e. how things are done), but also with the products (i.e. what gets done or used) and the producers: people or tools (i.e. who does it). Process, products, and producers (i.e. how, what and who) are therefore often considered the three major aspects of methodologies.

It is thus easy to see that “one size does not fit all”, that is, any given methodology cannot be optimal for every project or endeavour. Each project or endeavour will naturally have its own particular needs regarding how things must be done, what products must be used and generated, and what people and tools should be used. The characteristics of the organisation carrying out the project; the skills, culture and preferences of the people involved; the contextual constraints (such as time or resource limitations) are factors that shape what the methodology for a given endeavour should look like.

A methodology, to start with, should be adequate for the goals of the project or endeavour on which it is going to be enacted. That is, it should be able to produce the expected work products, from the available inputs, through the intervention of the appropriate individuals and tools, and using suitable techniques. This is self-evident and the most basic premise for a methodology.

In addition, a methodology must satisfy some needs related to its own sharing and communication. These needs arise in contexts such as that of collaborative work, when several individuals need to work in the same project or endeavour using a common, shared methodology; in cases like this, the agreed-upon methodology needs to be documented and circulated in a clear and unambiguous way so that everyone in the team applies the same
techniques, uses the same inputs, creates the same kinds of products and, in general, employs the same methodological framework for the common goal. This is especially important when the team is heterogeneous in composition, multidisciplinary and/or geographically distributed, and a shared language is not available or face-to-face communication is not feasible. Sharing and communication are also crucial in the context of the institutionalisation of methodologies; often, a government body, private company or other large organisation needs to establish a methodology (or a part of one) as a standard or “best practice” within a certain realm, be it the whole organisation or a subset of it (e.g. a particular department or a specific kind of projects). In order to successfully institutionalise a methodology, it needs to be properly documented and made available to the relevant stakeholders with minimal ambiguity; again, this is especially important when the parties involved belong to different disciplines and/or are geographically dispersed.

Finally, a methodology needs to improve over time, and the enhancements that result from this improvement incorporated into new projects and endeavours where the methodology is used. Usually, applying a methodology to a project means that the project is carried out with more or less success, but it does not necessarily entail a reflective process by which, in addition, the application of said methodology to the project is evaluated. This additional process, if performed, offers the opportunity to collect information about how well the methodology works on this particular project, what aspects of the methodology could be improved for future applications, and how.

In summary, a good methodology must be adequate, easy to share, and easy to improve.

In addition to these key properties, there is an extra factor that determines the success of a methodology, namely the amount of effort that is required in order to put it together. Methodologies are means to an end, rather than ends of themselves, and therefore they should be seen as part of the research or work infrastructure at our disposal. Any project team should better invest their time and resources in pursuing the project objectives rather than constructing a methodology for such a task; from this perspective, a methodology that is ready-made, or almost ready-made, is in principle superior to a methodology that must be painstakingly crafted by hand, perhaps reinventing the wheel in many aspects. This, however, clashes with the idea that “one size does not fit all” that we introduced at the beginning of this section: a ready-made methodology is necessarily of “one size”, so it would not be able to be adequate for every project or endeavour that could possibly arise. The solution to this dilemma lies in taking the best of each world: this is what method engineering proposes, as described in the next section.

3. Situational Method Engineering for Archaeology

Situational method engineering can be described as the discipline concerned with the engineering of methodologies according to each particular situation. The fact that it is the word “method” (rather than “methodology”) which appears in “situational method engineering” obeys to historical reasons, and the literature has repeatedly remarked that “method” and “methodology” should be understood as synonyms within this context (ISO/IEC 2007, section 3.3; Gonzalez-Perez and Henderson-Sellers 2008, 3). Very much like a civil engineer envisions, designs and creates roads, bridges and other infrastructures according to the requirements of their future users and the available time and resources, a method engineer envisions, designs and creates methodologies according to the requirements of their future users and the available time and resources (hence “situational”).
3.1 Overview of Situational Method Engineering

Situational method engineering appeared in the 1990s (Kumar and Welke 1992; Brinkkemper 1996; Rolland and Prakash 1996) as a sub-field of software engineering, due to the circumstance that software development projects were in much need of methodological guidance. However, it can be argued that situational method engineering is not intrinsically connected to software engineering, since its object of study is methodologies, whatever they are applied to; an example of this is the application of situational method engineering to the field of business processes described by (Gonzalez-Perez and Henderson-Sellers 2010).

Situational method engineering acknowledges that “one size does not fit all”, and that each methodology needs to be specifically situated, or adjusted to the project or endeavour to which it is going to be applied. At the same time, it avoids circumstances that involve reinventing the wheel by providing a solid knowledge reuse framework, so that methodologies are never created from scratch.

In particular, and from the perspective of situational method engineering, a methodology is not a monolithic entity, but an assembly of components that are carefully connected together after being selected from a pre-existing repository. Once a methodology has been created by assembling components, it can be enacted on an endeavour (i.e. applied to a project or other activity); during this process, the performance of each component can be assessed, and the result of this evaluation fed back into the repository in the form of improvements to the components stored there. This way, methodologies that are assembled in the future from the improved components take advantage from the accumulated enhancements that occur over time, thanks to the ongoing feedback loop (Fig. 1).
Interestingly, the terms “method” and “methodology” are apparently used in an interchangeable fashion in situational method engineering. These terms have very different meanings in theory of science, which roughly correspond to what one would expect from the “-ology” suffix in “methodology”: methodology is the discipline that studies methods. However, common usage of these terms, and a more descriptivist approach to their understanding, deviate from this; for example, the Oxford English Dictionary (Oxford University Press 2012) defines “method” as “a particular procedure for accomplishing or approaching something”, and “methodology” as “a system of methods used in a particular area of study or activity”; in other words, a methodology is a collection of related methods, rather than being the discipline that studies them. The situational method engineering literature has gone even beyond this, and made “method” and “methodology” practically equivalent; in fact, ISO/IEC 24744 (ISO/IEC 2007) explicitly defines the two terms as synonyms. For this reason, this paper will use them as such.

There are some additional aspects that need discussion. First of all, what is a method component? Components are often described in the literature along the lines of being reusable, atomic, self-contained packages of knowledge (Gonzalez-Perez and Henderson-Sellers 2008, 6). They are packages of information because they consist of text or similarly descriptive data that explains some relevant aspect of a methodology. They are self-contained because, in order to be understandable as true components, they must be self-explanatory, and as little reliant on other components or external sources of information as possible. They are atomic because they cannot be decomposed further in sub-components, and they are reusable because their major purpose is to be used in methodologies over and over again.

A second aspect to be considered is that the components that are selected from the repository and assembled together as a methodology are determined, to a large extent, by the methodological requirements of the endeavour to be performed. These requirements can be often described in terms of the products (or objectives) that the endeavour aims to achieve (such as documents, theoretical models or physical objects), the conditions of the environment where the endeavour will take place (such as time or resource constraints), and the sociotechnical characteristics of said organisational environment (such as the corporate culture towards innovation). OPFRO’s Process Consultant (Firesmith 2012) is an excellent web-based example of an implementation of methodological requirements, albeit applied to the field of software development rather than archaeology; however, the similarities are striking: methodologies are organised in phases and other stages, and roles such as “project manager” (the person who manages and monitors the project and its risks) are used throughout.

A third aspect that must be borne in mind, and which is depicted at the bottom of Figure 1, is that a metamodel is usually assumed to underpin the workings of a situational method engineering system. A metamodel is a formal language that can be used to describe the method components in the repository, the methodologies that are assembled from them, and the endeavours that are then enacted (Gonzalez-Perez and Henderson-Sellers 2008, 18). A metamodel, effectively, determines what types of components may be stored in the repository and then used as part of methodologies. A number of metamodels are available as official or de facto standards, and most of them cover the three areas that section 2 introduced as key for any methodology: the process, the products and the producers. For the remainder of this paper, we will adopt the ISO/IEC 24744 Metamodel for Development Methodologies (ISO/IEC 2007), an international standard that was created from the software and systems engineering fields, but which may be adapted to other fields.
(the next section explains some of the potential issues that may appear). ISO/IEC 24744 defines the following key concepts as part of any methodology:

- **Work units.** A work unit is a job performed, or intended to be performed, within an endeavour. There are different types of work units: *processes* (large-grained work units that operate within a given area of expertise, such as “Excavation” or “Pottery Analysis”), *tasks* (small-grained work units that focus on what must be done in order to achieve a given purpose, such as “Locate site on map”, “Wash and clean finds” or “Construct interpretative hypothesis”), and *techniques* (small-grained work units that focus on *how* the given purpose may be achieved, such as “Low-altitude photography”, “Probing” and “Surface collection of artefacts”; all of which may be applied to the task “Perform surface survey”).

- **Work products.** A work product is a thing of interest for the endeavour, either because it is used as an input to the process (such as an archaeological find), or because it is created as an interim or final result (such as an excavation report). There are different types of work products, the most relevant ones being: *documents* (durable depictions of a fragment of reality). ISO/IEC 24744, having been created from the perspective of software and systems engineering, does not provide for specific subtypes of work products that represent archaeological artefacts; this can be easily solved by extension, as described in the next section.

- **Producers.** A producer is an agent that has the responsibility to execute work units. There are different types of producers, the major ones being: *people* (individual human beings involved in the endeavour); *teams* (organized sets of producers that collectively focus on common work units, such as “the excavation team” of “the pottery lab analysts”); and *tools* (instruments that help other producers to execute their responsibilities in an automated way, such as “the balloon-mounted camera” or “the structured-light scanner”).

- **Stages.** A stage is a managed time frame within an endeavour. Work units only specify what is supposed to be done, but they do not say *when*; stages, on the contrary, establish the temporal framework for work units. There are different types of stages, the most relevant ones being: *phases* (managed intervals of time during which the endeavour changes levels of abstraction, such as “Formulate Project”, “Acquire Data” or “Interpret Data”); and *milestones* (managed points in time that mark a significant event within the endeavour, such as “Fieldwork Completed” or “Final Report Delivered to Client”).

ISO/IEC 24744 also establishes what attributes (i.e. properties) each of these concepts has; for example, the standard states that every *phase* has a start time, and *end time* and *duration*; and that every *document* has a *title*, a *creation time*, and a *status* (that can be either *initial*, *complete*, *accepted* or *approved*). Also, the standard defines what associations (i.e. relationships) exist between concepts; for example, it establishes that each *task* may cause effects on *work products*, and that these effects may be of several types (*create*, *modify*, *read only* or *delete*). It is beyond the scope of this paper to make a complete description of the semantics of each concept, attribute and association in ISO/IEC 24744; please see (ISO/IEC 2007) for more information.

As a last remark, it is important to realise that, so far, the concepts that have been introduced as part of ISO/IEC 24744 pertain to the realm of the endeavour rather than the methodology. In other words, work units, work products, producers and stages are “the stuff that the endeavours are made of”. A methodology, as we stated at the beginning of this paper, is a “specification of the process...
to follow together with the work products to be used and generated, plus the consideration of the people and tools involved, during an information-based domain development effort; that is, a specification of what potential endeavours may be enacted from it. It follows that if endeavours are composed of work units, work products, producers and stages, then methodologies must be composed of the specifications of these things. As a matter of fact, ISO/IEC 24744 defines a parallel set of concepts, suffixed with the “kind” word and organised in a similar structure, that pertain to the realm of methodologies: work unit kind, work product kind, producer kind and stage kind. A work unit kind, for example, is the specification (or “blueprint”) of a particular kind of work units that may occur in the endeavour. For example, if we were designing a methodology for a field project that is likely to use balloon-mounted cameras for low-altitude photography surveying, we may want to incorporate a technique kind named “Low-altitude photography” associated to the task kind “Perform surface survey”. These are method components that encapsulate reusable, atomic, self-contained knowledge, and as such are part of the methodology. The actual technique and task will only appear when the methodology is enacted on the actual endeavour (i.e. when we go to the field and start surveying), and correspond to each time that someone performs a surface survey (a task kind) by using low-altitude photography (a technique kind), on a particular date and time, in a particular place, and under particular conditions.

3.2 Applying Situational Method Engineering to the Humanities and Social Sciences

In principle, the application of situational method engineering to the humanities and social sciences in general, and to archaeology in particular, would result in the following benefits:

- Methodologies could be created quickly for the specific project or endeavour at hand, by assembling pre-existing components most of the time that are known to work in a wide range of situations.

- Methodologies would be easier to communicate and institutionalise, since the components they would be made of would be already documented, and the connections between them would be also known and explicit.

- Methodologies would be easier to improve over time, since enhancements to method components would be transferred to the repository, thus helping those who might use them in the future.

However, there are some potential issues that must be taken into account. Situational method engineering, as mentioned in section 3.1, emerged in the context of software engineering, and has been chiefly applied to software and systems engineering disciplines. It is, therefore, a discipline with a strong engineering view of the world, and this may produce problems of two kinds when applied to the humanities and social sciences. First of all, there may be technical issues caused by a mismatch in the epistemic approaches of engineering and humanities; where engineering is driven by a problem-solving attitude based on quantitative evidence, predictive models and underlying hypothetico-deductive methods, humanities and social sciences are often based on explanatory stances based on qualitative evidence, descriptive models and critical or otherwise non-hypothetico-deductive methods. In addition, engineering does not pay much attention to multivocality as a source of valuable knowledge, whereas the humanities and social sciences do.

Secondly, there may be additional problems due to the perception of these differences. Although we have not carried out a formal enquiry line in this regard, we have informally observed, and have anecdotal evidence of, a varying degree of uncompromising rejection
of engineering-mediated conceptualisations of the humanities and social sciences, especially from the field of anthropology.

Moving to a more technical level, another kind of issues that need to be taken into account, and which are strongly related to the first problem outlined in the previous paragraph, is that of metamodel suitability. ISO/IEC 24744, for example, has been designed from and for engineering, which manifests itself in a number of consequences:

- Methodologies are seen as productive methodologies, i.e. every methodology is supposed to create something from something else. Although this may be the case in engineering (given the dominant problem-solution paradigm), this may not be applicable to the humanities and social sciences, where constructing solutions to extant problems is often not the goal.

- The specific types of work products that are foreseen as part of the metamodel (models, documents, software items and hardware items) are clearly insufficient for the humanities and social sciences. New types of work products need to be incorporated.

- The complexity and diversity of actors (i.e. producers) and their effects on things (i.e. actions) that are typical of the subjects of study of archaeology and other fields of the humanities and social sciences are barely captured in the metamodel.

ISO/IEC 24744 does provide extension mechanisms (ISO/IEC 2007, section 9) that would allow for some adjustments, such as the addition of missing work product types. However, larger changes, such as the reconceptualisation of methodologies as not necessarily “productive”, would probably be impossible through metamodel extension, and would probably need the development of a new metamodel designed specifically with the humanities and social sciences in mind. This new metamodel would incorporate the management of multivocality as a key aspect of information (Gonzalez-Perez 2012), and would also take into account recent works in the characterisation of the cultural heritage domain such as (Gonzalez-Perez and Parcero Oubiña 2011).
4. Sample Scenario

For the sake of simplicity, let us assume a streamlined version of ISO/IEC 24744 with the structure depicted in Figure 2, using the ConML notation (Incipit 2012b; Gonzalez-Perez 2012). ConML is a simple modelling language that allows the representation of concepts, their properties and their relationships to other concepts.

Figure 2 (a metamodel) shows the types of method components that are allowed in our sample repository, and what relationships can exist between them. For example, there can be process (such as “Excavation” or “Surveying”), task (such as “Find artefacts” or “Locate sites”) and technique kinds (such as “Low-altitude photography”). Each process kind must be composed of one or more task kinds, and each task kind may be composed of no, one or more technique kinds. At the same time, there is a single subtype of stage with duration, namely, phase kinds (such as “Fieldwork” or “Lab Work”). Stages with duration are composed of process kinds. Task kinds, in addition, act upon work product kinds via action kinds. There are two subtypes of work product kinds: document kinds (such as “Project Report” or “Map”) and item kinds (such as “Archaeological Artefact”). The texts in the lower halves of the boxes describe attributes, i.e. what properties each concept may have. For example, all work unit kinds have a name and a description, and action kinds have a type (create, modify, read only or delete) and an optionality (mandatory, recommended, optional, discouraged or forbidden). A complete explanation of the semantics of the ConML notation is beyond the scope of this paper; please see (Incipit 2012b, 2012a) for additional information.

4.1 Sample Repository

A sample repository of method components can be described taking the metamodel in Figure 2 as a starting point. The following tables define each method component in the repository.

A number of things must be taken into account. First of all, the tables above show a subset of what a real-life repository would comprise. For example, Table 1 shows phase kinds that cover most of what many archaeological projects aim to do, including fieldwork, lab work, further interpretation and publication of results or outreach activities. However, Table 2 shows process kinds pertaining only to a few of these activities; none is illustrated, for example, in relation to the dissemination of results. This is necessary in this sample scenario for the sake of brevity, but a real-life repository would contain a much larger number of components, and of a much wider range of varieties.

In addition, and for the same reasons, the characterisation that the tables above give...
<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>In Process Kinds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find artefacts</td>
<td>Find artefacts in the target area, and package and label them in batches for later processing.</td>
<td>Excavation, Surveying</td>
</tr>
<tr>
<td>Clean artefact</td>
<td>Wash, brush and clean the artefact, as applicable, so that it is ready for description and long-term storage.</td>
<td>Excavation, Surveying, Restoration</td>
</tr>
<tr>
<td>Describe artefact</td>
<td>Perform a description of the artefact and enter it into the computer database.</td>
<td>Excavation, Surveying</td>
</tr>
<tr>
<td>Create replica</td>
<td>Create a replica of an artefact for research or exhibition purposes.</td>
<td>Restoration</td>
</tr>
<tr>
<td>Locate sites</td>
<td>Determine the locations of sites and mark them on the map in order to ensure ease of access later in the project.</td>
<td>Project Design, Surveying</td>
</tr>
<tr>
<td>Construct interpretative hypothesis</td>
<td>Build a hypothesis by using interpretation of the available archaeological and associated evidence.</td>
<td>Project Design, Excavation, Surveying, Restoration</td>
</tr>
</tbody>
</table>

**Table 3. Task kinds in the repository.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>For Task Kinds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-altitude photography</td>
<td>Taking pictures of the land’s surface through a camera attached to a balloon or other low-flying device.</td>
<td>Find artefacts, Locate sites</td>
</tr>
<tr>
<td>Surface collection of artefacts</td>
<td>Collecting artefacts found on the earth’s surface by intensively surveying the target area.</td>
<td>Find artefacts</td>
</tr>
<tr>
<td>GPS positioning</td>
<td>Using GPS devices to determine the coordinates of a point or area.</td>
<td>Locate sites, Describe artefacts</td>
</tr>
</tbody>
</table>

**Table 4. Technique kinds in the repository.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Action Kinds (ActionType, Task Kind, DeonticValue)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artefact Record</td>
<td>Created by Find artefacts [recommended], Created by Describe artefact [mandatory], Modified by Describe artefact [optional], Read Only by Construct interpretative hypothesis [recommended]</td>
</tr>
<tr>
<td>Map</td>
<td>Modified by Locate sites [recommended], Read Only by Find artefacts [recommended], Read Only by Construct interpretative hypothesis [recommended]</td>
</tr>
<tr>
<td>Project Report</td>
<td>Created by Construct interpretative hypothesis [recommended], Modified by Construct interpretative hypothesis [recommended]</td>
</tr>
</tbody>
</table>

**Table 5. Document kinds in the repository, plus their associated action kinds.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Action Kinds (ActionType, Task Kind, DeonticValue)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archaeological Artefact</td>
<td>Read Only by Find artefacts [mandatory], Read Only by Describe artefact [mandatory], Read Only by Construct interpretative hypothesis [optional]</td>
</tr>
<tr>
<td>Archaeological Replica</td>
<td>Created by Create replica [mandatory], Read Only by Describe artefact [optional], Read Only by Construct interpretative hypothesis [optional]</td>
</tr>
</tbody>
</table>
for method components is also artificially brief and condensed. A full-fledged repository would need longer and better descriptions for method components in order to make them self-explanatory, and would likely include additional attributes beside a name and a description.

Finally, a note is needed in relation to the associations between method components that are shown in the tables above. Associations in a repository are usually considered suggestions to the repository users rather than a prescriptive norm. For example, Table 3 states that the “Locate sites” task kind is suggested to be part of the “Project Design” and “Surveying” process kinds. This is mere advice that the repository encapsulates, derived from the accumulated experience of previous users. Any user is free to follow this guidance or, on the contrary, go against it by, for example, choosing to locate sites as part of the excavation processes, if that is meaningful to their particular methodological needs. The next section shows some exceptional choices like this in action.

4.2 Sample Methodology

Once a repository is in place, methodologies can be constructed by selecting the appropriate method components from it and assembling them together. The components to be selected are determined by the user’s methodological requirements, in terms of the methodology’s expected products, the properties of the projects where it is going to be used, and the organisational environment, as outlined in section 3.1. The way in which components are assembled together is determined, partially, by the associations that in the repository exist between them, and partially by the user’s specific needs. As explained in the previous section, the repository contains general advice on what the typical associations between components are, but specific methodological needs often require “rewiring connections” into customised arrangements.

Let us imagine that a finds lab at an archaeology museum decides to create a finds processing methodology for their own usage. The repository described in the previous section contains numerous method components that can be used straight away, but a few custom components also need to be introduced. Figure 3 uses the ISO/IEC 24744 graphical notation (ISO/IEC 2010) to show a lifecycle diagram of the methodology, i.e. one that focusses on its temporal perspective, depicting the phase kinds and what process “content” they have.

![Figure 3. Lifecycle diagram of the sample methodology, using the ISO/IEC 24744 notation. Two sequential phase kinds are shown, “Lab Work” and “Interpretation”. Some process kinds are shown for each phase kind.](image)

![Figure 4. Process diagrams of the sample methodology, using the ISO/IEC 24744 notation. For the sake of brevity, a very simplified view of what task kinds (depicted as ellipses) should be included in each process kind (depicted as “roundangles”) is given.](image)
Notice that the “Artefact Reception” and “Content Development” process kinds have been defined from scratch for this particular methodology, since they are not part of the repository. If they are shown to work well, the users of this methodology may decide to make them part of the repository for the benefit of future users.

Figure 4 shows process diagrams for the methodology; here, process kinds are displayed in more detail, so that their component task kinds are depicted. Process diagrams usually include the technique kinds that are associated to each task kind, and the producers that are responsible for the execution of each; however, the sample methodology is too simple to describe these aspects.

Finally, Figure 5 shows an action diagram describing how the task kinds of the “Restoration” process kind interact with the associated work product kinds: archaeological artefacts are modified by “clean artefact” tasks and used (i.e. “read only”) but not changed by “describe artefact” tasks, which in turn create artefact records. “Create replica” tasks, moreover, also use but do not change archaeological artefacts and optionally their records, in order to create archaeological replicas.

Similar action diagrams could be constructed to illustrate what other work products are affected by each group of task kinds, or what other task kinds affect other work products.

5. Conclusions

In this paper we put forward the hypothesis that situational method engineering can be applied to the construction, documentation and improvement of methodologies in archaeology and, possibly, the humanities and social sciences in general, as a response to a long-standing and increasing demand for attention to process-related issues within these communities. We have shown that situational method engineering is an excellent candidate to fulfil the methodological needs, and that, specifically, it would allow archaeologists to create methodologies that are extremely adjusted to each project at hand in very short times, since they are constructed from proven, pre-existing building blocks available through a repository.

In order to test this hypothesis and analytical claims, we believe that a proof of concept is a natural next step. Future actions in this direction include the analysis of existing methodological approaches in
Crafting Archaeological Methodologies
César Gonzalez-Perez and Charlotte Hug

archaeology in order to obtain “raw materials” for method components; the construction of experimental repositories; and actual testing of these repositories by assembling a few new methodologies and assessing their performance and advantages in real-life projects.

In addition, a deeper and perhaps more complex issue remains. As we pointed out in section 3.2, the application of a conceptualisation of archaeology that is mediated by an engineering approach constitutes a potential barrier to the adoption of method engineering in this field. Qualitative research needs to be carried out in order to assess the real impact of this problem.

References


Boundary Concepts For Studying the Built Environment. 
A Framework of Socio-Spatial Reasoning for Identifying and Operationalising Comparative Analytical Units in GIS

Benjamin Vis
University of Leeds, UK

Abstract:
The layout of built environments is a rich source of spatial data. If we are to understand the significance of the physical presence of these spatial properties to the social development of the residing society we need to move beyond a purely empirical representation and quantitative analysis of acquired spatial data. This paper presents both a theoretical conceptualisation and a conceptualisation of built environment layout data that brings boundaries as the elementary constitutive socio-spatial property to the fore. On the basis of this conceptual understanding of boundaries in our live world and in spatial data an ontology of analytical units can be devised, which are abstract, yet empirically identifiable. They form the basis of an emergent comparative methodology of restructuring spatial datasets which will permit the interpretive use of GIS analyses of built environment configurations in the future.

Keywords:
Socio-spatial Theory, Comparative Methodology, Boundaries, Built Environment, Spatial Analysis

1. Introduction

The use of computation in archaeology forces us to think in a quantifiable way despite our end aims often being qualitative, that is, the interpretation and better understanding of our data in their context. When spatial data are concerned a GIS is typically developed to store and organise data by their spatial component or reference. Next to that ability, most GIS software also features opportunities to carry out spatial analyses. These are usually based on standardised statistical calculations on the basis of the spatial, locational and metric, attributes of our data and any associated information. In order to carry out such analyses one’s data needs to be structured according to a format that is recognisable and compatible with the technique. But even when this is achieved, there is no necessary reason why the standard GIS analyses are at all appropriate for the research aims. Because of the empirical and specifically quantitative nature of GIS, using it to interpretive ends requires a careful consideration of both data structure and analytical suitability. Such endeavour forces the active questioning of the assumptions behind the analytical technique and the nature of the data operationalised by it.

Spatial data is virtually omnipresent in archaeology, as material remains are first retrieved in a location and spatial context. The focus here is limited to spatial data on built environment layout, one of the richest spatial datasets available to us. The aim is to provide a theoretical conceptualisation that leads directly to methodological opportunities for developing GIS analyses that are of interpretive use. In this effort this paper concurs with Gillings’ (2012) proposition that GIS practitioners should develop their own theories and subscribes to the potential of reconciling critical and quantitative research approaches (see Kwan and Schwanen 2009) by presenting the development of a socio-spatial theoretical ontology that is GIS ready.

This paper will discuss an approach for the comparative interpretive study of the built environment that the author has been developing

Corresponding author: B.N.Vis10@leeds.ac.uk
initially as a general theoretical treatise (Vis 2009) and subsequently as research aimed at capturing the resultant elementary property of the built environment layout as it develops (Vis forthcoming). This elementary property is established as consisting of boundaries, and allows for the analysis and social constitutive interpretation of spatial configuration. The approach is developed from an interest in the significance of the physical presence of built environment layout for social development.

Having first constructed much of the theory required to address this interest, this paper presents the conceptual basis for the development of a methodology for the study of appropriate data on the built environment, to which the following underlying questions are central: how could we study the spatial information derived from urban built environment configurations in order to understand its significance for the way society develops? Or more specifically, how does spatial layout play a role in the development of society and how does spatial layout accommodate its residing society? After a brief recapitulation of the enabling and informing conceptual series that form the theoretical framework, this paper will present a conceptualisation of spatial layout data. The theory and data conceptualisation together will form the foundation of intellectual understanding and empirical identifiability on which resultant analytical units can be defined. The intention with these analytical units is that they may be applied to spatial data that can in turn be used in GIS analyses. In conclusion, this paper will briefly present the prospect of two major computational challenges following from the conceptualisations. By overcoming the computational challenges, it is hoped that in the future this will enable the full potential of the emergent methodology and reveal interpretive opportunities.

2. Theory

The study of the significance of the physical presence of the built environment for social development requires a constitutive understanding of what the built environment is and does. That is to say, how is the built environment a result of human action from the bottom-up? Which are the elements out of which the built environment emerges in continuous formative processes? And, what is the restrictive and enabling role of the built environment in the social life-world? As has been argued elsewhere (Vis 2009; Vis forthcoming) these concerns directly touch upon the nature of our being: both human being in the spatial world and human being in the social world.

Departing from fundamental Heideggerian existential philosophy, we need a concept of human being based on how this ‘being’ is conditioned. Firstly it should be recognised that we are temporally and physically conditioned by our bodies. Ideas about how we are ‘embodied’ have received great attention in anthropology (Bourdieu 1977; Csordas 1990; Ingold 2000; Low and Lawrence-Zúñiga 2006) in attempts to overcome the mind-body dichotomy. Through this process the body has become established as a site of lived experience and embodied agency (Joyce 2005). However, this does not yet fully account for Heidegger’s being-in-the-world and the phenomenological bi-implication of man and environment (see Ingold 2000) in which the body is not only physically capacitated, but of which the body is intrinsically part. Howes (2005) has proposed the concept of ‘emplacement’ to express the immediacy of the sensuous interrelationship of mind-body-environment, which holds that the environment is both physical and social. Griffiths and Quick (2005) make an appeal to study spatial configurations following a similar premise of human being in a physical and social environment. It is this holistic advancement concerning the situatedness of the condition of human being, moving beyond the concept of embodiment, which is taken forward here.
foundations of time-geography (Hägerstrand 1975; 1976; Pred 1977)). Our bodies allow us to interact with the surrounding environment that is not our body and in those actions construct relations through which we get to know the world experientially (cf. Schütz’s (1967) constitutive phenomenology). Our being is acting, and each action has a time and place, while it incorporates the physicality that allows us to transform the physical properties of the world we live in (cf. Pred 1984; Pred 1986). Each action expresses an ordinal decision of preference: we act with the subjective expectation that we improve our situation (Von Mises 1998; Vis 2010). Ordinal decisions do not adhere to normative rationality. Our actions give our being time-space specific particularities. Our resultant understanding of the social and physical environment, is self-referential (Schütz 1967; Von Mises 1998) and our understanding of others regardless of language can therefore never be complete (Vis 2009). The biographies we acquire by perceiving the relationships we construct in interactions (Pred 1984; Pred 1986) with our environment play a constitutive role in how we subsequently act (Griffiths and Quick 2005; Vis 2009). Society as a specific emergent bundle of inter-personal, social, relations develops from a constant merger of the axes of human time, human action and human space (Vis 2009). The physical spaces that are created in this process result from our perceptive being-in-the-world and our emplaced or experiential knowledge of the causative actions (see Ingold (2008a) on inhabiting the world). Even though, in our being, our environment is both immediately social and spatial, it is heuristically advantageous to start an understanding of the built environment from theoretical conceptualisations of being in the spatial world and being in the social world separately.

Vis (forthcoming) demonstrates how the conceptual series of being in the spatial world and being in the social world indicate a specific datum of the built environment. This datum is the boundary, the physical difference between one thing and another, based on the principal importance of differentiation for inhabiting the world. Boundaries deserve further conceptualisation if they are to become the analytical unit of research for those seeking to understand the constitutive socio-spatial significance of the physical presence of the spatial configurations we make and inhabit ourselves. To explain how one arrives at boundaries as the essential datum of the built environment and how these should subsequently be conceptualised, the conceptual series of Vis (forthcoming) will be briefly summarised below.

It is helpful to start thinking about the perceptive being in the spatial world by imagining the space of the world as empty, i.e. the absence of difference. Although the world cannot technically be empty if a human inhabitant is present, we can imagine the experience of an abstract empty world as unintelligible. There is nothing within and nothing confining the space we are in, meaning nothing to engage in interaction with and nothing to relate to. A comparison can be made to Gibson’s (1979) absence of textural difference and Ingold’s (2008a) empty void.

As Gibson (1979) notes, however, our world is ‘furnished’ with all kinds of affordings features. Again, disregarding that any human inhabitant would render this impossible, we can imagine an abstraction of our world as if (cf. Ingold 2008a) no-one was ever there, but with all the natural, uncultivated features that the world we know and inhabit has (cf. Deleuze 1984). If we were to inhabit such a world it would be inhabitable because of the intelligibility of the affordances or opportunities for survival provided by the already extant physical properties of this world (cf. Appleton’s (1975) prospect-refuge theory). This is what will be called the primordial world, i.e. before human inhabitation. This abstract notion excludes the effects of human presence in space, while retaining our essential ability to perceive differences in physical properties of the world
and experience the opportunities associated with them.

In contrast, by experiencing the spatial world as equalitarian we encounter the unintelligibility of emptiness again. Equalitarian space is perceiving the spatial world as continuously indistinctive by repetitions of the same, stretching beyond the horizon. This is both an abstract and concrete notion as, though unlikely to be completely accurate upon closer inspection, being on a perceptively limitless sandy plain with a smooth overcast sky (cf. Gibson 1979; Ingold 2008a), or a dense pine forest would leave us uncomfortably inapt to direct our actions. Even humanly created environments might resemble this: e.g. suburbs of uniform blocks of flats. It is in our capacity to interact with and encounter the physical properties of the world that we might discover their uniqueness and affordances as well as our ability to modify these properties.

When we start acting upon the physical properties of the world we inhabit, we may leave involuntary and deliberate traces or marks of our lived experience of the world. Marking space is the introduction of lasting physical effects to properties of the processes of formation (see Ingold 2008a on fluxes) that are already ongoing in the world we encounter. This creates or enhances (contextually) the perceptible distinctions of physical properties, which improves the intelligibility of the spatial world. These distinctions can be as diverse as the physical properties that can be perceived and our technical abilities to modify materials. It is the start of a process by which we ‘fill’ space previously unaffected by human presence; the introduction of human processes into the world by acting (participating) in that spatial formation. Recognising such distinctions is a continuous subjective self-referential process: the perception and experience resulting in conceptual constructs which become anchored in our knowledge (see Schütz 1967).

The conceptualisation of being in the spatial world is concluded by the process of filling space, which consists of the complete transformation of physical properties of the spatial world. It is an extension of marking and leaving traces, as it essentially links up those modifications to come to complete containment and division of the spatial world in partitions by acts of physical transformation. The spaces of our being in the spatial world have now become their own distinctions, built by our own actions, i.e. the built environment. It is in creating spaces, architectonic structures according to Van der Laan (1983), that we make the spatial world habitable, containing the prerequisite of the intelligibility of differentiations. Table 1 presents a comprehensive overview of the conceptualisation of being in the spatial world. When everyday life takes place in an environment of completely filled space, we approach a likeness to the intensity of inhabiting the world that is urban.

The existence of an environment of completely filled space highlights the absence of other human beings from this conceptualisation. The physical capacity of a single human being to transform the world is limited. These other human beings also inhabit the spatial world and therefore are equally part of our environment and our lived experience thereof. As the same space can never be occupied twice at the same time, each human being has an individual, unique, position and situation throughout its life path (cf. Pred 1984; 1986). Consequently our knowledge and understanding of inhabiting the world is never equal to someone else. Nevertheless, the commonality of our species and the sharing of the live world enable an empathetic self-referential understanding of others to various degrees of inaccuracy (see Vis 2009).

Being in the social world is initially composed of encounters. Encountering by the logic of uniqueness of position and situation requires space, no matter the scale of the encounter. Encounters can also occur indirectly,
by recognising the marks or transformations left by others, so it is not necessary to have two people present simultaneously in their mutually perceptive environments (Vis 2009). When we encounter and interact we negotiate personal territories by setting the interpersonal distance for the encounter we are comfortable with (see Hall’s (1959; 1968) proxemics). This distance setting is a first step in spatially organising the social world.

Building on Giddens’ ideas, the flexible notion of projects, setting a goal for an activity either individually or institutionally with the participation of others (Pred 1981; 1984; 1986), accounts for the emergence of systems. The constant convergence of life paths and associated biographies of experiential knowledge through participating in projects creates social interaction spaces in a constant process of becoming. In this light a communal sense of place emerges. Following Pred (1984; 1986) these projects incorporate an environment which can be physically affected. A start is made with a socio-spatial world.

The emergent systems are not confined to any scale and are autopoietic, auto self-creative (see Koch 2005). They are understood to simultaneously constitute themselves through differentiation from within and from without, and the constitutive influence of the social and spatial (physical) environment is mutually dependent. Systems produce understandings of space through a process of differentiating themselves in their environment, making the environment socio-spatially intelligible. This differentiation is not simple opposition, as a distinction towards the environment also constitutes the system. Existing spatial properties form constitutive environments, while individual participants’ biographies of lived experience form a social context for the actions that let systems emerge (see Bruun and Langlais 2003; Vis 2009). Because the physical properties of the environment are structurally coupled to inhabitation and projects let interrelational systems emerge, the differentiation of socio-spatial systems is necessarily related to the subdivision of filled space.

At this stage the notion of the built environment becomes ‘populated’ by the human being inhabiting their environment. The built environment thus consists of physical distinctions that are approximations of the negotiated outcomes of the differentiation between socio-spatial systems. As the scale of these systems is flexible they may adhere to various spatial extents of built distinctions, while the built distinctions themselves influence how interaction between participants of all operating socio-spatial systems is mediated, accommodating the residing society in formation.

3. Understanding Boundaries

The descriptive term ‘built distinctions’ used so far solely refers to physical effects and therefore is unable to capture the full complexity of how we understand and have conceptualised the outcomes of differentiation. A built environment composed of built distinctions is a built environment of bounded spaces. Therefore the more neutral and flexible

<table>
<thead>
<tr>
<th>Spatial differentiation</th>
<th>Conceptual nature</th>
<th>Intellectual character</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty space</td>
<td>Abstract</td>
<td>Unintelligible</td>
</tr>
<tr>
<td>Primordial space</td>
<td>Abstract</td>
<td>Inhabitable</td>
</tr>
<tr>
<td>Equalitarian space</td>
<td>Abstract/concrete</td>
<td>Unintelligible</td>
</tr>
<tr>
<td>Marked space</td>
<td>Concrete</td>
<td>Made intelligible</td>
</tr>
<tr>
<td>Filled space</td>
<td>Concrete</td>
<td>Made habitable</td>
</tr>
</tbody>
</table>

Table 1. The conceptualisation of the differentiation of being in the spatial world.
term ‘boundaries’ is henceforth used, as it can be invested with both the ideational and empirical relevance of differentiation. It should be noted that much attention has been given in sociological (e.g. Lamont and Molnár 2002) and anthropological research (e.g. Pellow 1996) to what could be called socio-cultural, political or formalising and administering symbolic boundaries as constituents of social categories. Such metaphorical considerations only receive an implied place within the argument presented here. As will transpire, the primary concern is with working towards empirical application on the basis of physical difference and therefore boundaries here do not explicitly include differences in class, nation, kin or cultural identity, though it is acknowledged these are a relevant influence on how the role of physical differentiation is understood. As a case in point Lawrence (1996, 33) stated that: “It was commonly at the border between private and collective spaces (by the entrance door or at the windows) that residents engaged in expressive behavior with kith and kin.” This quotation illustrates Lawrence’s strong assertion that boundary thinking is able to converge many disparate social and cultural research interests.

As demonstrated in Lawrence (1996) here the intricate relationships between boundaries as conceived and understood differences and boundaries as perceived and experienced physical differences are explored on a level that can be empirically operationalised. The constitutive theory above has established that boundaries are the elementary property of the built environment which conveys the socio-spatial significance of its physical presence as its residing society develops. In order to work with the notion of boundaries as an intellectual, social, understanding and an empirical, social, reality an ontology of respective analytical units needs to be developed (see Vis forthcoming). This ontology of analytical units should respect the notion that “[boundaries] are not merely artifacts of the differentiation of places. They are intrinsically related to the actions they contain.” (Rotenberg 1996, 56).

Boundaries merit the attention of an entire field of studies, mostly in sociology and human geography, but also philosophy. As will be demonstrated abstractions made in select writings within boundary studies have been very beneficial for making explicit the conceptualisation of the direct relationship between the theory above and spatial data on built environment layouts. Spatial layout data result from documenting and measuring physical differences. It is essential that we fully understand what our data represent and purports in the light of the current research aims, before we can develop analytical units that respect the empirical reality they are identified in.

A focus on boundaries rather than the entities we are prone to think with (see Jones 2009; 2010) is exactly what proponents of boundary studies have argued. Jones (2009) argues that the heterogenesis produced by ‘deterritorialisation’ reveals a socio-spatial complexity (sometimes inferred as multiplicity) that is normally disguised by categorical divisions. In the same way the built environment can be seen to disguise socio-spatial complexity in its fixity and apparent ‘container units’, i.e. entities or categories. In the light of the constitutive theory above, Abbott’s (1995, 860) assessment of the temporal priority of boundaries over entities falls into place: “social entities come into existence when social actors tie social boundaries together in certain ways. Boundaries come first, then entities.” Jones (2009, 175) argues that “the problem is not the categories themselves, but, rather, the way the boundaries around the categories are cognitively understood as closed and fixed even when we know intellectually that they are open and fluid. Consequently [...] the key process is the bounding and delimiting of the categories used to understand the world.”

It is clear that proceeding from this vantage point means accepting boundaries as inchoate processes of making sense of the world (Jones 2009). Boundaries are sites of difference and
rather than starting from entities, we should start investigating entities with the boundaries that create them: ‘things of boundaries’ in Abbott’s (1995) words. The differentiations boundaries make are the results of processes of negotiating and renegotiating, fixing and re-fixing distinction in the same way as autopoietic socio-spatial systems assert themselves.

Spatial data on the built environment, however, are devoid of this dynamism, representing aspects of a live world as static. This static can be regarded as a stage of fixity in the inchoate processes. The notion of stages in development is also important for Abbott (1995) who explains the structural resilience of emergent entities through their defensibility in several dimensions of difference. In the socio-spatial world, boundaries offer distinctions of territoriality, intelligibility, habitability, and familiarity. Boundaries emerge from the repeated performance of systems internally defining them to their outside in the context of the constitutive environment and all participating biographies. A materialised boundary resulting from such processes is naturally persistent (resilient). Often there is no necessity for the currently performed socio-spatial system to match the original emergence of the materialised boundary, as many activities are to an extent spatially independent (see Sayer 2000). Boundaries emergent from systemic performance are thus not necessarily related to function, but rather to the accommodation of interaction opportunities. A static representation therefore approximates a stage of consolidation (see Vis 2009) in the development of the built environment. Such persistence is a degree of inertia in physical construction (building) and disguises the formative processes in which they are ephemeral (i.e. can be transformed). They allow us to see the built environment as constructed of edges (sites of physical difference).

This is where we start considering the nature of our spatial data once more. As Schaffter, Fall and Debarbieux (2010, 260) rightfully note: “[b]oundary studies, however broad and theoretical, must [...] remain alert to spatiality and materiality, and not just to processes of construction.” The concepts that were introduced to ‘reconstruct’ the processes of the constitution of the built environment reveal to us, intellectually, the socio-spatial complexity the built environment represents and the perceived entities it is composed of conceal. This offers us a starting point for the conceptualisation of analytical units on the basis of boundaries as the constitutive elements of built environment configurations. Understanding the socio-spatial complexity of boundaries before devising analytical units reverses the impoverished understanding resulting from subscription to the oversimplicity of the static binary character of categories. It has been demonstrated how society and space cannot be separated (see Vis 2009). Though constructed by human beings themselves, the physical existence of boundaries becomes a constitutive part of the experiential biographies of human beings encountering and inhabiting the physical world. The built environment thus consists of measures of persistence of affordance, which result from the coherence of elements as entities or emergent aggregates. The temporal priority of boundaries thus gives way to an expression of mereological causality (see Smith and Varzi 1997; 2000). In the words of Smith (2001, 3): “[o]bjects and processes can each be conceived as being put together or assembled out of (respectively: spatial and temporal) proper parts.”

The ontology we desire for studying built environment layout depends on identifying those elements of it that mark the differentiating negotiation of different socio-spatial systems, i.e. the elements that link up constitutively to form the physical entities we perceive. The aim is not to categorise these entities, but to reveal and understand the relative coherence of the diverse and fuzzy relational differences they consist of.
4. Built Environment Configuration

Although the above provides a thorough understanding of the complexity of boundaries, this protean character also renders the term boundary elusive. Their intellectual understanding as a site of difference is not a constant and the word can refer to the physicality of lines, edges, borders, barriers, divisions, etc. In order to better understand the nature of built environment layout data, the distinction between *fiat* and *bona fide* boundaries as proposed by Smith and Varzi (1997; 2000; Smith 2001) will be used.

Essentially the opposing definitions of *fiat* and *bona fide* boundaries are remarkably simple, but their application has profound effects on the understanding of spatial data. *Bona fide* boundaries are those distinctions that are based on spatial discontinuity (e.g. holes, fissures, slits) or qualitative (physical) heterogeneity (e.g. material constitution, texture, electric charge). *Fiat* boundaries are those distinctions that are based on differentiation without association with spatial discontinuity or qualitative (physical) heterogeneity. This completes the opposition (see Fig. 1 for a schematic representation). The definitions apply equally to inner and outer boundaries, which in theory does not limit their use to any distinction (Smith and Varzi 1997; 2000; Smith 2001). It may seem inaccurate to introduce a difference between inner and outer boundaries, because how to define *a priori* what can be regarded as the entities allowing for inner boundaries? Once put in a processive perspective it becomes acceptable that when an entity is established (e.g. a hut) inner distinctions of its parts could be made (e.g. various zones).

In addition, both *fiat* and *bona fide* boundaries may form entities, or, in Smith and Varzi’s (1997; 2000) terms: objects. *Bona fide* objects could be a body, ball or cheese, whereas *fiat* object could be a property, a hemisphere or the North Sea. For the latter, the North Sea, it is apparent that along the coastal lines *fiat* and *bona fide* boundaries coincide. One should add though that the lines delimiting the North Sea on maps are entirely *fiat*, because although one may argue it represents the physical distinction between land and sea, we know that the tides are in constant movement. Therefore, no coastal line can ever be truly represented by a static line. It should be noted that most *fiat* boundaries do not exclusively depend on human *fiat*, but (as phenomenology suggests) involves the underlying material also. *Bona fide* objects cannot also depend on *fiat* boundaries (the examples in the preceding are an adaptation of Smith and Varzi (2000)).

Smith and Varzi (2000) further distinguish between individual *fiat* boundaries, and *social* *fiat* boundaries. The first pertains to the arbitrary choice made on the basis of individual perception: typically resulting from a single act at a given time. The arbitrary choice of setting a *social* *fiat* boundary depends on the perception of all participating human beings. *Social* *fiat* boundaries can also be imposed abstractly and may appear relatively isolated from local causal change (e.g. policy driven distinctions). Finally, *bona fide* objects are necessarily all connected (in the continuum of the material world), whereas *fiat* objects may be scattered. Yet, some of the scattered *fiat* objects may be unified in *fiat* objects of a higher order: e.g. island groups (Smith and Varzi 2000). These
additional details are useful to note in relation to the preceding theory. Nevertheless, above all fiat and bona fide boundaries essentially refer to the presence or absence of a precise physical distinction.

As bona fide boundaries are material, i.e. they have divisible bulk, they necessarily occupy space. Bona fide boundaries must be part of the bona fide objects they circumscribe. This means that bona fide objects are ineluctably the owners of their own boundaries, whereas the environment they are embedded in is open. Bona fide objects thus have open complements. The exception are negative objects: where a void occurs in the surrounding material surface the resulting bona fide object is defined from the outside by the bona fide boundary of its ‘host’. The beginning of the object then partly depends on the virtuality of fiat (Smith and Varzi 2000; Smith 2001). As mentioned before, according to the time-geographical principles of Hägerstrand (1975; 1976) the same spatial location cannot be occupied twice simultaneously. This necessarily means that bona fide boundaries cannot coincide, as also acknowledged by Smith and Varzi (2000). However, “fiat boundaries, because they are not possessed of divisible bulk, do not occupy (fill out) the space where they are located; hence they can be perfectly co-located one with another.” (Smith and Varzi 2000, 416) These occupational differences are pivotal for understanding how the concepts devising our built environment data come about.

From this follows that bona fide boundaries may become and coincide with fiat boundaries and fiat boundaries may give way to additional (modified) fiat boundaries. However, the only way a fiat boundary can become a bona fide boundary is through the according transformation of the physical properties of our environment, i.e. materialisation. From that moment onwards they enter a dialectic relationship in which the fiat boundary and the bona fide boundary continually become one another. That is, the processes of formation are inchoate. The bona fide boundary becomes an immediate social reality in space that is reacted to in the same way as the preceding environment. That means the ideational (human or social) boundary causes the empirical (physical) boundary and the empirical in turn influences the ideational and so on.

What these oscillations between fiat and bona fide understandings of boundaries mean for our understanding and preparation of spatial data on built environment layouts, which

<table>
<thead>
<tr>
<th>Built environment data type</th>
<th>Nature</th>
<th>Acquired as</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1   Built boundaries</td>
<td>Bona Fide</td>
<td>Measurements and documentation of physical transformations of the environment, i.e. materialised differentiation</td>
<td>Unprocessed empirical observations (data as acquired)</td>
</tr>
<tr>
<td>2   Boundary lines</td>
<td>Fiat</td>
<td>A construction and reduction to the outlines of the full, continuous extent of materialised differentiations</td>
<td>Processed and simplified version of empirical observations</td>
</tr>
<tr>
<td>3   Boundary visualisation</td>
<td>Fiat</td>
<td>A geographical representation of the outlines</td>
<td>Integral interpretive geographical organisation in an outline plan (full configuration)</td>
</tr>
<tr>
<td>4   Boundary line types</td>
<td>Capture a fiat understanding of bona fide</td>
<td>An application of an ontology of analytical units on the outline plan (boundary line type ascription)</td>
<td>Dissected plan of geographically anchored analytical units</td>
</tr>
</tbody>
</table>

Table 2. The breakdown of built environment layout data expressed according to the principles of bona fide and fiat boundaries.
<table>
<thead>
<tr>
<th>Boundary line type</th>
<th>Empirically identifiable principle</th>
<th>Social relation to interaction opportunities</th>
<th>Exemplary indication (for western or globalised cities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Closing boundaries</td>
<td>Operates on the basis of seclusion of a continuous spatial arrangement from the surrounding configuration with the material property that the boundary can be closed off towards its outside, making it a dominant. It is also a solid (i.e. no internal arrangement of outlines)</td>
<td>Interaction opportunities are quite stringently internalised as distinct from the outside, though there is a mutual (in)direct orientation between the solid dominant and the surrounding configuration</td>
<td>These boundaries typically circumscribe buildings of any sort or size</td>
</tr>
<tr>
<td>2 Facing boundaries</td>
<td>Operates on the principle of the orientation for soliciting interaction from the surrounding configuration</td>
<td>Is the site of solicitation of interaction with a dominant</td>
<td>These boundaries represent the doorways or entrance ways into a building</td>
</tr>
<tr>
<td>3 Associative boundaries</td>
<td>Operates on the basis of dependence on a single dominant it is directly associated with and, in a conjunction including possible other associative boundaries, forms an adjoining configurative complex with</td>
<td>Interaction opportunities are mediated between the openness of the surrounding configuration and the related dominant</td>
<td>These boundaries are typically associated with gardens or any plots and surfaces belonging to a specific building</td>
</tr>
<tr>
<td>4 Extended facing boundaries</td>
<td>Operates on the principle of orientation in uninterrupted connection to a facing boundary by dependence on any boundary associated with a dominant</td>
<td>Is the site of indirect solicitation of interaction with a dominant, proceeding is no necessity</td>
<td>These boundaries are typically associated with garden gates or courtyard entrances, etc.</td>
</tr>
<tr>
<td>5 Directing boundaries</td>
<td>Operates on the basis that it directs interaction along opportunities for further boundary crossings in parallels</td>
<td>Interaction opportunities are directed along the boundary crossings that constitute its sides, connecting all sorts</td>
<td>These boundaries are associated with the street network, access and pathways</td>
</tr>
<tr>
<td>6 Disclosing boundaries</td>
<td>Operates on the basis of guiding interaction towards opportunities for further boundary crossings in multiple directions rather than a single particular direction with necessary (in)direct connections to solid dominants</td>
<td>Interaction opportunities are freely organised, yet directed in multiple directions which in several cases will eventually lead to soliciting interaction with solid dominants</td>
<td>These boundaries are associated with square-like spaces in well integrated urban situations with several associated buildings</td>
</tr>
<tr>
<td>7 Enclosing boundaries</td>
<td>Operates on the basis of seclusion from the surrounding configuration with the material property that the boundary can be closed off towards its outside, making it a dominant while containing solid dominants</td>
<td>Interaction opportunities are restricted by solicitation between the openness of the integration within the boundary configuration and the configuration with solid dominants that it circumscribes</td>
<td>These boundaries are typically associated with city walls and gated communities</td>
</tr>
<tr>
<td>8 Mutual boundaries</td>
<td>Operates on the principle that it is simultaneously associated with or encompassing a distinct subset of several solid dominants</td>
<td>Interaction opportunities are indirectly directed to several solid dominants and mediated between the openness of thoroughfare</td>
<td>These boundaries are associated with a specific group of buildings without any preference as to which it provides access, such as shared porches, cul-de-sacs and communal space in gated communities</td>
</tr>
</tbody>
</table>
Boundary Concepts For Studying the Built Environment. A Framework of Socio-Spatial Reasoning
Benjamin Vis

<table>
<thead>
<tr>
<th></th>
<th>Boundary Type</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Opening boundaries</td>
<td>Operates on the principle that it creates open, accessible connections towards its outside, while being an integrated part of the configuration</td>
<td>Interaction opportunities are freely organised, with no prerequisites for boundary contexts and the possibility of thoroughfare</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>These boundaries can be described as park-like spaces, e.g. garden plots, urban fallow, parking surfaces</td>
</tr>
<tr>
<td>10</td>
<td>Neutral boundaries</td>
<td>Operates on the principle of neutrality, which results from ambiguity and the absence of singular associations, can occur in virtually any context</td>
<td>Due to the absence of a unambiguous relation to a residing socio-spatial system, crossing the boundary creates no difference from the surrounding non-dominant configuration</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>These boundaries tend to be the left over areas in less optimally used built environment configurations and also some delimited functional areas connected to streets (e.g. electricity supply)</td>
</tr>
<tr>
<td>11</td>
<td>Man-made boundaries of unoccupiability</td>
<td>Operates on the basis of negativity, can occur in most contexts</td>
<td>Negativity means there is no residing socio-spatial system, in this case because an area cannot be occupied by human beings</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Structures that create a unoccupiable surface area, such as ponds, canals, architectural talus, narrow gaps, etc.</td>
</tr>
<tr>
<td>12</td>
<td>Not man-made boundaries of unoccupiability</td>
<td>Operates on the basis of negativity, can occur in most contexts</td>
<td>Negativity means there is no residing socio-spatial system, in this case because an area cannot be occupied by human beings</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Steep slopes, natural bodies of water, etc., which are contained in the built environment</td>
</tr>
<tr>
<td>13</td>
<td>Not man-made negative boundaries</td>
<td>Operates on the basis of negativity, can occur in most contexts</td>
<td>Negativity means there is no residing socio-spatial system, in this case because it marks the end of the built environment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>'Nature' or not fully cultivated areas</td>
</tr>
</tbody>
</table>

Table 3. The current iteration of the boundary line type ontology.

have become conceptualised as compositions of materialised differentiations, is presented in Table 2. The table comprises four stages in which spatial data becomes converted into boundary conceptualisations: from the initial acquisition of spatial information, through to processing that information into meaningful ascriptions from which analytical units come forth.

A few things should be noted following the data breakdown in Table 2. Firstly, the methodology that is emerging here offers an idealistic presentation of data acquisition. In practice it could logically be expected that researchers will use already extant maps and plans, both commercially and academically produced as well as potentially historical map documents. The advantage is that legacy data can easily be used. It should be noted, however, that extant data and plans require treatment with meticulous care. For example, how to discern primary entrances to single buildings? The scope of this paper prohibits a full exposure of the methods for using specific extant data and plans.

Secondly, the production of or reduction to outlines conceals their underlying intricate consideration. Outlines should follow from adherence to the single (the smallest scale) continuous residing socio-spatial systems. As the internal arrangements and ongoings of such systems cannot be known from the theory, the internal arrangement or interior design of space(s) within outlines should not be represented. In practice the definitions of the analytical units (below) will help at this stage to make informed decisions, though judgment calls cannot be avoided. As bona fide objects contain their own boundaries, built boundaries with an occupiable mass will be regarded as internal arrangements.
Finally, Table 2 does not reveal the assumed temporal congruency of the spatial data. That is, the data utilised should represent a single moment in time and will therefore depend on historical and archaeological determination of this. Even though it is recognised that the nature of any plan and data acquisition prohibits true simultaneity, once a plan is produced it should be regarded as an atomic snapshot. Atomic here means an immediate and inseparable, in other words: no time passes, everything occurs at once. Only this assumption will allow full analytical operationalisation, as the dataset needs to be both complete and represent features that are known. The implications are that special effort should be made to make sure all features mapped did indeed exist contemporaneously (avoiding anachronisms), while also conjectures may be necessary to fill in any indicated blanks in the spatial continuum.

5. Conceptualisation of an Ontology of Analytical Units

From Table 2 it has already become apparent that the analytical units uniting the intellectual understanding and empirical identification, which are dependent on boundary lines, will be called boundary line types. A critical reader might remark that analytical units and types mark a retreat into immediate entities and categories again. This paradox can be overcome by noting that in observing the empirical reality of built boundaries, we assume and ascribe a fixed nature to that reality. The empirical recognition of analytical units based on the data derived from built boundaries, depends on that fixed representation of the empirical reality. Abbott (1995) states that boundaries as sites of difference should be regarded as atomic units (points or elements rather than categories) to which differences can be attributed. By extension of his argument boundary line types are the smallest meaningful elements of the boundary lines composing the built environment configuration. Boundary line types, then, are the atomic units to which differences in meaning can be attributed in terms of their theorised constitutive socio-spatial significance.

As boundary line type ascription comprises an integral redescription of the built environment configuration, the formulation of boundary line types needs to be ontological (i.e. cover all boundary lines). Kropf (2009) asserted that a good usable ontology for comparative research should consist of a consistent, coherent and comprehensive set of definitions, which are general enough to be used comparatively in all thinkable contexts, yet specific enough to identify them as analytical units in the empirical reality of datasets. Conform Kropf’s reasoning the definitions of the boundary line types together will mereologically resemble readily perceived composite entities and analyses could mereologically generate aggregates conveying the complexity of the larger scale.

Both the preceding theoretical conceptualisation and the resultant boundary line type ontology (Table 3) are compliant with critical realist thought and methods (see Sayer 1981; 2000; Lawson and Staeheli 1990; Pratt 1995; Yeung 1997). Critical realism is a philosophy of science that serves as a corrective of the more popular, improper relativist contradictions of post-modernism (Pratt 1995; Sayer 2000; see Wallace (2011) for a commentary on archaeological theory). Critical realism is especially helpful when bridging the divide between interpretive and empirical science and addressing ontological concerns in a controlled fashion based on causation. This research is an example of exactly such effort, although as can be recognised throughout, aspects of post-structuralist and phenomenological thought have been found highly informative within this framework.

The ontology presented in Table 3 has resulted from constant contestation of initial a priori concepts in a process called iterative abstraction. This is one of the key conceptual methods proposed in critical realism. This
means that several iterations of the boundary line types presented below have preceded the current list. These previous iterations were deemed flawed in recursive application to various datasets. Although the current set of conceptualisations appears relatively stable, it cannot be precluded that (data of) empirical conditions and experiences exist which will require a reappropriation of one or several definitions, including possible mergers of existing types or additions of entirely new types. Nevertheless, the definitions should be neutral and specific enough for application in contexts regardless of culture area or time period. Imperfections and flaws in the practical adequacy of the current conceptual ontology can only be revealed through critical utilisation.

It is impossible to present the full definitions and associations with the levels of socio-spatial significance for the interpretation of analytical outcomes and coherence patterns in a single paper. The definitions need to take into account several factors that are often difficult to express, such as their empirical identifiability, which is strongly dependent on spatial context, their relation to some basic material properties of impermeability, and, not least, the effectual role they play in relation to the accommodation of interaction opportunities they create, which depends on the side of the boundary one is located at (seen from the inside or the outside). This is simply too elaborate to include here, which means Table 3 is necessarily a brief shorthand indication of the actual definition. Though the process of boundary ascriptions will simply cover all boundary lines of a given configuration, it should be noted that no boundary line receives a single definition. Depending on their context of adjacent features each boundary line is part of a boundary line type identification at least twice.

Finally, the ontology is strictly speaking incomplete as the relations between boundary line types are not defined. It is an ontology of types, which means their mereology is not predetermined and therefore the way they become related as parts of the redescription of the built environment configuration is fully dependent on the time-space specific context. The boundary line types are elements in a redescription of the spaces that compose the built environment in a way that respects and resembles the processes of bordering featuring in Deleuze and Guattari’s ‘geophilosophy’, despite its materialist nature often used in metaphorical applications (Woodward and Jones 2005), but rather concretised in a process of empirical identification. Any analysis will allow interpretation following the theoretical framework of constitutive socio-spatial significance of patterns that fully emerge from the specific conditions of any site at any time. Indeed it is this final interpretation of the outcomes of analyses on the basis of theoretically informed units, a formal ontology that is regarded as interpretation in computation (Leszczynski 2009), that defines this use of GIS firmly as being of qualitative purpose (Kwan and Schwanen 2009). By restricting the socio-spatial theory to the empirical reality of inhabiting the world and the empirical reality of our own data, regarding the social as a holistic realm of interaction opportunities, a true comparative understanding of built environments is enabled appropriate to the aims of the research interests formulated at the outset.

6. Conclusions

The work presented in this paper is part of a Ph.D. research in progress. It has argued strongly for an approach to studying built environment layouts using conceptualisations of boundaries resulting from a socio-spatial theory of human beings inhabiting the world, and creating a society. It is hoped that a truly comparative methodology can be constructed from this approach, for which this is the basis.

In way of a conclusion, it is most informative to indicate the two major computational challenges that will need to
be addressed to enable the full potential of the conceptualisations presented in this paper. Both are currently being assessed as operationalisations in ArcGIS software. The first challenge is to enable the topological analysis of the boundary line types, once digitally applied to the outline plan. The principal step is to access the topological information contained in the compiled boundary line type identifications, after which the development of analytical computing tasks can start. Both stages, however, require the building of original software for ArcGIS. The second challenge is to maintain data integrity through time-slices, i.e. snapshot plans, which is a requirement for topological analysis through time. Unfortunately the standard error margins of georeferencing and georectification prevent the simple achievement of topological integrity over time. As the theoretical conceptualisations are constitutive, it is expected that the most important advances of the resultant methodology will entail revealing and studying processes of built environment development. This technical challenge is related to the development of historical and temporal GIS, which at the moment still primarily relies on correlating maps representing specific points in time. During a recent workshop (Mapping Spatial Relations, their perceptions and dynamics: the city today and in the past) at the University of Erfurt participants noted that an informed liberal approach to the interpretation of maps can be acceptable historical practice. This could well be the way forward to conjecture historical continuity in the built environment through time, mitigating the loss of accuracy in time depth.

Continuing research will explicate all of the methodological implications the presented conceptualisations hold. Many of the practical issues will be specific to the context of the built environment selected and the data that is available. By creating comparative concepts and a general methodology it is expected that the analytical power of GIS software will become of interpretive use in generating socially constitutive understandings of built environment layouts beyond the limitations of mere correlation.

Notes

1. GIS relies on computation, making it necessarily quantitative, and measured or quantified data entries, which render it empirical. In recent years concerns have been raised on the positivist and reductionist perception this causes amongst qualitative researchers and how to advance the qualitative or critical non-empirical use of GIS (e.g Kwan and Schwanen 2009). It is acknowledged that the quantitative nature of GIS, despite its limitations (see Leszczynski 2009), does not withstand its use for qualitative purposes and approaches that are sensitive to societal complexity, diversity, and becoming. In archaeology McEwan and Millican (2012), Gillings (2012), and Hacigüzeller (2012), all within the space of a single, year have published on the need and opportunities to push GIS approaches further with proper theorising and ontologies, phenomenological sensitivity, understandings of affordance, and non-representational thought. This paper does not serve to contribute to this specific polemic, but rather offers the first steps of progress made by following such approach for the purpose of the social study of the built environment.

2. In archaeology the original metaphysical (Husserl) and existential (Merleau-Ponty) branches of phenomenology have grown to become much more popular than the here prioritised sociologically inspired constitutive phenomenology of Alfred Schütz. A full discussion of how these branches relate can be found in Campbell (1981). The rationale for preferring Schütz’s phenomenology results from the strong action theoretical connection, which is central to the theoretical arguments here.
Furthermore, his cogent proposition of a social subject with social experience in an individual life-world, resonates better with the social scientific concerns of this research. It allows a ready connection to a social constitutive or emergent perspective of society in systems theoretical sociology (cf. Giddens 1984), while simultaneously maintaining a connection to the experiential reality of inhabiting a physical world (see also Vis 2009).

3. It is acknowledged that emptiness can be a philosophically laden term, but it is felt that argument is helped here by appealing to the commonsensical imagination of empty space.

4. Affordance is a still contested ecological concept (Jones 2003). Broadly speaking in this paper affordance features still close to Gibson’s (1979) original propositions: the opportunities for and consequences of acting upon the physical properties of the world. This is separate from perception of objects’ substance and surface properties, because it concerns the action centred implications of their presence. The idea and lived experience of emplacement (Howes 2005) does bring these two positions close together. For an operationalised view of the concept of affordance see the discussions on materiality by Knappett (2004; 2007) and Ingold (2007; 2008a; 2008b) and related ideas discussed in Webmoor and Witmore (2008) in material culture studies.

5. Despite the static wording, these differentiations are in constant formation and therefore fluctuate as the system is performed by its participants. This may occur in relation to physical differences that were created as part of the projects letting the system emerge. This physical difference tends to be a more persistent marker than the operating differentiation.

6. Note that the emergent complex of the built environment results from purposeful, intentional acts, but that the exact outcomes of actions cannot be foreseen nor can any subsequent acts of others. The aggregate of the complex in its entirety therefore is always partially unintentional, i.e. not the congruent product of any single human being’s intention.

7. Archer (2005) gives a useful summary of main strands of scholarly literature that have influenced thinking on built space from socio-cultural, economical and political perspectives. It should be noted that the social understanding that is key here is rudimentary and does not distinguish between time-space specific realms of the social. Rather it is restricted to seeing the social as a holistic realm of interaction opportunities which are constituent of society and coherence within it (cf. Giddens 1984).

8. It should be noted that nowhere in this research a boundary indicates geo-political and other administrative categories such as districts, parishes and national borders.

9. The full boundary line type definitions will be presented as part of a prospective Ph.D. thesis by the author.

Acknowledgements

The author would like to thank Andrew Evans, David Bell and Penelope Goodman for their relentless support of this research and their comments on various iterations of the conceptualisation presented here, and Siân Horan Smith for constructive comments and debate in preparation and during CAA 2012. A special thank you is due to the University of Leeds for funding this research with a University Research Scholarship.
References


Everything Flows: Computational Approaches To Fluid Landscapes

Dimitrij Mlekuž
University of Ljubljana, Slovenia

Abstract:
Current computational approaches to mobility in archaeology are based on deterministic models and assumption about optimal behaviour of past agents. Movement, especially when viewed over very long time, exhibits patterns that can be described as flows. The patterns of long-term movement are constrained by the terrain but are also formed by the interaction of agents with the traces left by agents who have moved in the same landscape in the past. Tools presented in the paper, based on cellular automata, offer a new way of approaching past movement that complements existing deterministic models such as optimal paths.

Keywords:
Flow, Movement, Landscape, Morphogenesis, Cellular Automata

1. Introduction

People, animals, things and substances do not move around randomly. Their movement, especially when viewed over very long time, exhibits patterns that can be described as flows. Archaeological record is full of material traces that channelled those flows and were made by them: paths, hollow-ways, roads, streets ... (Edgeworth 2011a; 2011b).

Flows are spatial, temporal, and above all, material. Flows move people, animals, things, substances... into new positional and relational contexts with other things and create new material encounters, allowing new and different flows to emerge (Sheller and Urry 2006; Hannam et al. 2006; Urry 2007; Aldred and Sekedat 2011a). Flow is liquid; it is constantly re-shaping itself, often in surprising ways.

The landscape is thus constituted as a fabric of movements and flows that themselves produce their own forms. Places are defined by their relations with other places, by flows that connect them. Once flow stops, the places as such disappear. Movement is thus continuously generative, and if we want to understand it archaeologically, we need to develop approaches that can cope with the fluid conditions of generative practices (Aldred and Sekedat 2011a; 2011b).

Despite this, movement is usually represented in static form, as image, and is usually limited to the mapping the material traces of flows. This is problematic, as it does not acknowledge the emergent aspects of flows. In the paper we are going to present computational tools, developed to cope with the fluidity of flows, exploring their emergent properties in space and time.

2. Topography of Connectivity

Past movement is usually modelled using cost surfaces. A cost surface is a continuous surface, in which each part of the landscape (raster cell) is assigned a cost, or “price” to reach that point from source point. Cost is computed from distance and “friction” or difficulty of moving across the raster cell expressed either as time or caloric expenditure. Least cost pathways between the source of cost surface and target are traced by using some kind of optimising algorithm.
But cost-surface analyses can have much more general analytical use in archaeology. The cost-surface, where the cost is expressed as time, represents the maximal possible extent that an agent can reach in given time (or time budget). The binary map that defines geographic locations that an agent can occupy with given time budget may be called potential path area.

Potential path areas are the most general way of describing movement. They focus only on constraints that operate upon the moving actor. In this way they show where movement could occur, given the constraints. They are extremely imprecise, but also very accurate.

The idea of potential path area can be expanded to include every point in the landscape (Mlekuž in press). The potential path field is thus a sum of potential path areas within a given time budget from every location in the landscape. It is constructed by summing the potential path areas with allocated time budget from every raster cell of the study area (Fig. 1).

$$PPF = \sum PPA_{ij}$$

**Figure 1.** Calculation of topography of connectivity. Where $PPA_{ij}$ is potential path area from location $i,j$ and $PPF$, potential path field or topography of connectivity which is calculated as the sum of potential path areas from each location.

The result is a continuous raster surface, where value of each cell provides information from how many points within a given time budget is connected. Cells with higher values are connected to more locations within time budgets (Mlekuž in press).

Potential path fields are at first difficult to interpret. They do not describe specific acts of movement, or movement from specific location. They provide a new idea of landscape, based on its terms of connection to other places in landscape. We can also call it “topography of connectivity”. Instead of seeing landscape as a set of delimited places, topography of connectivity allows us to approach it as a set of connections between places. In this way, places are not defined by their location but in terms of their interconnectivity with other places.

The topography of connectivity identifies areas that are better connected and thus shows the structuring potential of the landscape itself on the movement and practices in landscape. Changing the time budget the relation between time and structuring effect of the landscape can be explored (Fig. 2).
Topography of connectivity is highly dependent on time budget. With limited time budget, absolute values are generally small, as one can access only limited number of locations form each spot. Local differences thus have huge effect on topography of connectivity. Thus differences between close locations are sharply pronounced. Choosing small time budget results in fine-grained differences between locations. As time budget increases, topography of connectivity becomes absolutely higher, but more and more and more smoothed out and less grainy as close locations become more similar in terms of connectivity. We can observe the emergence of large-scale patterns, such as basins, barriers and corridors.

Thus, in addition to more busy areas in the landscape, we should also expect to find non-places, back-areas and barriers. We can measure how connected are parts of the landscape, where interaction is facilitated by the landscape, where flows can occur and where they are blocked.

The topography of connectivity can be analysed using morphometric parameters such as slope, aspect (or orientation), and morphometric features (Fig. 3). Thus the slope of topography of connectivity gives us the change in speed. On steep areas travel slows down or speeds up, depending if one is traveling down or up on the topography of connectivity. Steeper the slope, larger is the change in speed. ‘Flat’ areas indicate areas where speed of movement is constant. Aspect or orientation of the topography of connectivity is a direction, where the change in connectivity is largest.

Using morphometric tools topography of connectivity can be subdivided into features such as pits, peaks, channels, ridges, passes and planes. These features must be understood in terms of rates of change of three orthogonal components of potential path field and can be interpreted as channels, funnels or barriers. By applying the prominence filter (Llobera 2001) the difference of connectivity between the point and other points in the landscape within the arbitrary neighbourhood can be approached. In this way we can isolate parts of the landscape within the neighbourhood that are more accessible within the specific radius, thus locations that are better connected than their locations inside the neighbourhood. We can see how topography of connectivity is then structured into network of well-connected locations and islands of backwater locations (Fig. 3).

All these analyses makes possible to explore how landscape is structured into islands of backwater and corridors, channels of movements and flow on different temporal and spatial scales. All these more or less distinct features, places or locales may be important for the understanding how people, animals and stuff moved around the landscape and how they perform activities. Such places and features can function as the corridors, which channel the flow of movement, structure daily practices and where activities can occur.

3. Morphogenesis of Flows

Landscape is not a passive background, but a flow itself, constituted through fabric of movements and flows that themselves produce their own forms. Sanford Kwinter brilliantly presents this idea when he describes the rock climbing: rock is an actualisation of time and
space that uses the fluctuating conditions of movement, intuition and invention to assemble:

The mineral shelf represents a flow whose timescale is nearly unfathomable from the scale of duration represented by the electrolytic and metabolic processes of muscle and nerves — but even at this timescale ... singularities abound ... This very rock face ... now swarms with individualized points, inhomogeneities, trajectories, complex relations ... the climber's task is less to ‘master’ in the macho, form-imposing sense than to forge a morphogenetic figure in time, ...to become soft and fluid himself... and to engage the universe's wild and free unfolding through the morphogenetic capacities of the singularity. (Kwinter 2001, 31)

Landscape comes into being through actualization of flows. But topography of connectivity encompasses all possible connections. It is charged with potential movement, but this movement is never actualised, it is, as Gilles Deleuze (1991) describes it, “virtual”. How are these virtual flows in the landscape of connectivity realised, actualised? What is the relation between the possible, manifested in landscape of connectivity and the actual emergence of flows?

In relation between actual and virtual, the actual does not resemble the virtual, as something preformed or pre-existing. The relation of virtual to actual is not one of resemblance but rather of difference, innovation, and creation. Actualisation does not programmatically reproduces what is already here, formed and given in advance. The actualisation is invention, an active dynamic process of differentiation and evolution. This is what is called morphogenesis.

The morphogenesis can thus be defined as a form obtained as a result of processes of differentiation and evolution. The opposite of emergence is design: it is a form conceived by a designer, which will be used as a blueprint for its realisation.

We are therefore interested in the morphogenesis of the flows in the landscape, their actualisation rather than properties of the topography of connectivity itself. We will approach the morphogenesis using cellular automata (CA).

CA are discrete dynamical systems, composed of lattice of cells, each of which may be in a predetermined number of discrete states. Cell states are updated based on states of its neighbours and its own state in the previous time step. Time in the model is also discrete; on each time step, every cell updates its state using a transition rule that takes as input the states of all the cells in its neighbourhood (Ilachinski 2001).

A famous example of CA is Conway’s “game of life”. Game of life consists of (theoretically) infinite lattice of square cells, where each cell can either be alive or dead. After each time step, the cells are updated based on few simple rules: a dead cell with exactly three live neighbours becomes a live cell, a live cell with two or three live neighbours stays alive, where in all other cases, a cell dies or remains dead. These simple rules produce surprisingly complex patterns, such as oscillating patterns, patterns that move around the lattice, patterns that produce other patterns and self-replicating patterns.

CA is therefore ideal tool for studying the morphogenesis, the emergence or pattern (Deutsch and Dormann 2001). Patterns are not designs imposed on matte; instead they emerge, evolve and grow out of local interactions. It is extraordinarily difficult to predict, a priori, how a CA will evolve from an arbitrary starting-point; usually the only way to do it is to work the calculation through explicitly. As such, CA are often used to study pattern formation in biological systems, such as the growth of microorganisms, dynamics of cellular tissue and tumors, and formation of pigment cell patterns (Deutsch and Dormann 2001).
4. Model of Morphogenesis of Flows

We designed a CA to model the formation of flows on the topography of connectivity. CA model is based on a model of pedestrian dynamics that takes its inspiration from the process of chemotaxis. Chemotaxis is a phenomenon where some insect insects create a chemical trace to guide other individuals to food places. This is also the central idea of the active-walker models used for the simulation of trail formation (Schadschneider et al. 2009).

CA is composed of lattice of cells where agents are situated. Multiple occupations of cells are allowed. Agents interact with their immediate Moore neighbourhood (eight surrounding cells), and can move to any of neighbour cell, based on transition probabilities (Fig. 4). The transition probabilities for all agents depend on the strength of different “floor fields” in their neighbourhood in such way, that transition in the direction of larger fields is preferred (Eq. 1).

\[
S_{ij} = Ne^{KpP_{ij}} e^{KtT_{ij}} D_{ij} \quad i, j \in \{-1, 0, 1\}
\]

(1)

Transition probabilities of agents in the cellular automaton where \( S_{ij} \) is transition probability to the cell \( i,j \); \( P_{ij} \) is topography of connectivity field at location \( i,j \); \( T_{ij} \) is traces field at location \( i,j \) and \( D_{ij} \) is direction field at the location \( i,j \); \( K_p, K_t, N \) are constants.

We distinguish between three kinds of floor fields. First one is topography of connectivity, it is static, computed in advance and describe the property of landscape, namely its connectivity. All other thing being equal agents will move in the higher connected cell within the neighbourhood.

The dynamic floor field consists of traces created through the motion of agents. The agents create trace by entering the cell. The main purpose of this field is to transform effect of long range interactions following the agents some distance ahead, or in the past into a local interaction, so that agent interact with the trace and not the agents themselves. At the beginning it is zero, but gets increased whenever agent enters the cell. Furthermore, it has its own dynamics, through diffusion and decay. And there is a direction field, which defines the general, or preferred direction of movement.

Behaviour of agents is based on local interactions between them and fields. Each cell can be empty or occupied by one or more agents. The update is done in parallel for all agents.

In the first experiment half of the cells in the landscape was randomly populated with one agent. A destination cell in upper left corner was selected as destination. The model was then used to observe the patterns agents produce as they move towards the destination.

We can observe how quickly order emerges in the form of lanes of agents moving towards destination point. The random pattern at the beginning of experiment quickly evolves into several columns of agents reaching the destination. This is also reflected in the traces field, where dendritic pattern of traces quickly emerge. Traces coalesce into several massive
flows that guide most of the agents towards the destination. Thus most of the agents ultimately reach the destination by following one of three major traces (Fig. 5).

In the second experiment, agents populate the discrete area and then move to the destination cell on the other side of the lattice. Two major flows of agents emerge, which channel the most of the agents towards the destination cell (Fig. 6).

Nearest neighbour interactions exhibit many of the collective effect and self-organisation phenomena such as lane formation, bottlenecks.... The form of flows cannot be conceived in advance, it is not planned, but emerges from simple local interactions between agents and landscape. It is a morphogenetic figure, an actualization of virtual movement through unfolding of the topography of connectivity.

5. Conclusions

Current computational approaches to mobility in archaeology are based on deterministic models and — implicitly or explicitly — invoke the theoretical assumptions about rational behaviour of agents.

We approached the emergence of flows as a dynamic morphogenetic figure. The patterns of long term movement are constrained by the terrain but are also formed through the interaction with the traces in the landscape, left by agents who have moved in the same landscape in the past. As such they present a new computational approach to the past movement that complements widespread deterministic methods.

We want to stress that our tools are not producing non-problematic representations of flows, as an end points that record and communicate information about movement. Even more, we believe that those representations would ultimately be misleading and wrong. Instead, these tools should be seen as explorative activities, that do much more than resemble, they take the place of the original situation and thus enable us to intertwine our practice of modelling with the past flows and movement patterns (cf. Hacguzeller 2012 on problems of representation in GIS).
Study of morphogenesis of complex flow patterns is crucial to future mobilities research as it intersects with scientific research into dynamical systems. It also suggests that when we study landscape, movement, we need to move beyond predefined, planned network topologies, and consider fluid, surprising, emergent topologies (Hannam et al. 2006; Sheller and Urry 2006; Urry 2007).

References


Reliability of the Representation of a Distribution: a Case Study on Middle Bronze Age Metal Finds in the Seine Valley

Estelle Gauthier
UMR CNRS 6249, Chrono-Environment, France

Maréva Gabillot
UMR CNRS 6298, ARTeHIS, France

Abstract:
Controlling the uncertainties and the documentary lacks of an archaeological corpus should be the priority to interpret correctly the data distribution. The conception of a reliability map in order to estimate the quality of the inventory, the creation of a representation map so as to highlight the over- and under-representations and finally their combination, the confidence map, are efficient tools to spot out the zones concerning which there is a risk for a defective interpretation. We propose here, an application on a case-study of Middle Bronze Age metal hoards and isolated finds, for which the only really determining factor is the possibility for discoveries since they came almost exclusively from chance finds. It can be evaluated after the contemporary population density, and the modern works implying the digging up the soil.

Keywords:
Reliability of Data Distribution, Documentary Lacks, Chance Finds, Metal Hoards, Isolated Finds, Bronze Age, Seine Valley, France

1. Introduction

The Research carried out by Workgroup 3 of the ArchaeDyn II programme focuses on spatial dynamics with respect to the production, circulation and consumption of raw materials and manufactured products over the long term (Gauthier et al. 2008; Gauthier et al. in press).

In this aim it uses 13 datasets created for the needs of other research programmes. They concern various raw materials or finished objects, dating mainly from the Neolithic to the Iron Age, in different regions. This implies a high heterogeneity of data and it brings up the necessity to control the state of knowledge of the studied areas. Can we say that the inventory is sufficiently reliable to realize a spatial study of the circulation and consumption of products? Data distributions can reflect biases in the archaeological documentation caused by the varied conditions of the research. In order that the team shall be able to use these data and correctly interpret the results of the spatial analyses, it is firstly necessary to estimate the reliability of the distributions.

Here, we present a case-study on the Middle Bronze Age metal hoards and isolated finds in the lower and middle Seine valley. This corpus is particularly interesting for the elaboration of a method for the estimation of the distribution’s reliability because the data essentially come from fortuitous discoveries. So the corpus and the repartition of the data are above all influenced by the finds made at the occasion of agricultural activities, development work, exploitation of natural resources (such as dredging, mining), etc. We propose to examine the impact of the factor of the “possibilities of discovery” on the archaeological data and we are going to see whether, in the present state of the research, it is possible to control it.

3 The proposed model does not attempt to take into account each ‘chance event’ linked to local and punctual activities, but simply to estimate the global capacity of each municipality to have provided finds in order to understand the reliability of the distribution maps.
The “reliability map” is the most useful tool for achieving this aim (Oštir et al. 2008). By knowing the conditions for data recording in each part of the study area, it is possible to qualify the distribution of the inventory on 3 levels of reliability. The criteria we have defined are the quality and quantity of the archaeological research (presence of specialists, quantity of excavations and field surveys, quality and exhaustiveness of the research, accessibility and visibility of the site), the conservation of the sites (destruction, erosion, built over) and of the artefacts (loss, dispersion in private collections), the circumstances of discovery (agriculture, roadwork, dredging and quarries, other development works...), the accessibility of the data (in situ, publications...) and the attention paid to the sector (Gauthier et al. 2012, 67-68). Finally this “reliability map” has to be combined with a “representation map”, showing under- and over-representation of the data in the distribution, in order to create a “confidence map”, the most complete tool for pointing out biases in archaeological research and their impact on the distribution of the data (Saligny et al. 2008, 25-44).

2. Presentation of the Case Study

The database relating to the Middle Bronze Age (from 15th c. to 14th c. B.C.) metal finds in the lower and middle Seine valley comprises 153 hoards and 524 isolated finds, that is to say 677 discoveries. The particularity of this example is that the bronze hoards and the isolated metal finds are usually chance finds. By definition a hoard is a set of objects gathered together and buried or immersed, out of a domestic or funerary context (Verger 1992). So it is a closed context with very limited dimensions hence it has lower chance to be discovered than a settlement.

This region has a highly unequal data distribution. We observe large areas of concentration along the Seine, especially around Rouen, Paris and Compiègne, while others are devoid of discoveries (Fig. 1).

The “Archaeological research” factor has little influence in this case (Fig. 2) because this kind of discoveries is very rarely found during planned excavations or field surveys (the number of field surveys with metal detector is negligible for this dataset essentially composed of discoveries from the 19th and early 20th centuries). This inventory was constituted by an exhaustive bibliographical research (Gabillot 2003). Bronze hoards – those which were reported to archaeologists – were seldom left unnoticed and are generally published. A part of the discovered bronzes is always misappropriated before reaching the archaeologists who will never know these finds (remelted, privatized or sold on the illegal markets etc.), so this part cannot neither be estimated nor be taken into account. The discoveries made at the same time as the excavation or the field survey of another site are rare; we find sometimes hoards or isolated bronzes in megalithic contexts, but their proportion is negligible. So the impact of archaeological activity on the discoveries of...
Bronze Age hoards and isolated finds can be generally considered as insignificant.

Hence, the “circumstances of discovery” is the only one criterion that really influences the distribution of these data (Fig. 2). Looking at the recorded contexts of discovery (Fig. 3), two-thirds is unknown, but a large proportion of the known contexts are related to rivers (dredging) and many to agricultural activities, development work and material extraction. Nevertheless it is clear that the river finds are more often mentioned and over-represented.

3. Protocol of Analysis

So the most important criterion that has to be taken into evidence is the possibilities for discovery. As the data were localized per municipality (the localization of the majority of these non recent finds is based on municipality names, so it cannot be specified more punctually), we must work with the polygons of the present administrative boundaries, we should not summarize information into a dot as one might have the reflex to do so as an easy way out.

Therefore our objective here is to estimate the potential of each municipality for having provided discoveries. The potential of each municipality is estimated independently of archaeological discoveries and then in a second step, it is compared with the density of known finds.

Municipalities that have had the highest chance to furnish discoveries are those with the highest density of modern human occupation and where human activities implying digging up the surface are the most intense: urban regions, zones with important development works such as roads, highways or railways, places of material extraction, dredging and cultivated lands. Conversely forests, moors, meadows, marshes, lakes have a low potential. We used contemporary data, which limit the punctuality of this analysis, since the situation may have changed since the 19th and the beginning of the 20th centuries.

3.1 The Choice of the Criteria

The criteria taken into account are treated one by one. In a first step, the proportion of the surface of each municipality covered by the considered zone type; then a score system

---

4 We underline that it is not a predictive modelling.
5 Corine Land Cover or IGN (French National Institute of Geographic and Forest Information) data.
is established in order to compare the different municipalities. For populated areas and areas of human activity, the following 3 steps were realized: a calculation of the cumulated area per municipality, the calculation of its proportion to the total area of the municipality and the classification of the values into 4 categories (Fig. 4).

The same method is used on other surface data: agricultural zones and for forests (Fig. 5). For roads (Fig. 6), as well as for the railroads and for main navigable dredged rivers, which are linear data, the following 2 steps were realized: a calculation of the cumulated length per municipality and the classification of the values into 4 categories. For places of material extraction (mines, sand quarries), we considered their presence or absence on the municipal territory.

In the next step, the previous results are compared with the archaeological data. Can we actually perceive an impact of each factor on the finds? Of course for any discoveries there should evidently a possibility for discoveries, but does the quantity of discoveries increase as the level of the possibility for discoveries rises? A comparison to a Chi-square calculation is an appropriate method to answer these questions.

The density of hoards and isolated finds per municipality has been classified into 5 categories (0: absence; 1: low; 2: moderate; 3: high and 4: exceptionally high). A Chi-square calculation makes it possible to estimate the theoretic number of municipalities in case of independence between the density and the factor into consideration (Groupe Chadule 1994, 86-87). Thus, the comparison with the real number reveals over-representation or under-representation in the different classes.

When the low values of density of vestiges are over-represented in case the considered factor is slightly represented on the municipal territory, while high values of density are under-represented in case the considered factor is highly represented on the municipal territory, we can suppose a positive correlation between the two parameters: the density of vestiges rises with the frequency of the zone type in consideration. This is what we can expect for the zones the most dug up.

Contrary, when the low values of density are under-represented in case the considered factor is slightly represented on the municipal territory.
territory, while the high values of density are over-represented in relation with the low values for the factor in consideration, we can came to the conclusion of a negative correlation; that is what we expect for the slightly altered zones which offered few possibilities for discoveries.

The different factors in consideration are crossed one-by-one with the density classes of the discoveries (Fig. 7). The following five criteria influence the number of discoveries: places of material extraction, dredging, inhabited zones, roads and railways. The more they are present the most important is number of discoveries in the territory of the municipality.

97% of the municipalities where slightly modified and unanthropized zones are predominant, did not yielded or a very few Middle Bronze Age hoards or isolated finds, that is a logical result and reinforce the precedent results.

Contrary the cultivated zones and the forests have given unexpected results. For cultivated lands to which the finds should correlate well, the absence of correlation may be caused by a great number of municipalities concerned as well as by the less systematic character of the works (the digging up was shallower). We can also suppose that the finds were perhaps less frequently declared (in addition the mechanized recent agriculture restricts the possibilities of observation). For municipalities with large forested areas, in which the discoveries should be under-represented, it is more difficult to explain the opposite results. But without the municipalities having a high value for the precedent factors, the results are close to the independence. This means that these two factors (forests and agriculture) are less influential.

As the zones the most dug up and the most altered (it is to say those having high potential after the 5 factors influencing the discoveries that were defined above) are the most likely to yield discoveries, they can be considered as the most reliable.

3.2 Creation of a Reliability Map

It is possible then to make a reliability map through evaluating the global potential for chance finds. It supposes a combination of the selected factors and the establishing a scoring system in order to qualify the potential of each municipality. The level of the 5 factors influencing the discoveries serves as a basis for the calculation of the scores, but the “secondary” criteria (agriculture, forests, inhabited and not overbuilt zones) are taken into account as well at a lesser degree in order to readjust these scores. Indeed, even if the discoveries do not present a particular correlation, the slightly altered zones give little chance for the discoveries, while the repetitive digging up the soil during the agricultural activities could have
been the occasion for the discovery of metal hoards or isolated finds.

Therefore each municipality receives a score to evaluate its potential for discoveries. Since these values are scattered so they are reclassified with the aim of establishing a 4 level reliability map (Fig. 8): very low reliability (municipalities of very low potential), low reliability (municipalities of very low to average potential), moderate reliability (municipalities of good potential) and high reliability (municipalities of very high potential).

3.3 Validation of the Defined Scoring System and the Reliability Map

Two tests were then realized to validate the scores for potential for discoveries and the corresponding reliability levels:

Fig. 9 presents the average value attributed to the potential of the municipalities classed after the density of the known hoards and isolated finds. More the municipality yielded the vestiges the higher is the average score: the municipalities, which did not provide discoveries receive few points while the municipalities providing a lot of discoveries receive the highest scores.

The second test is the comparison, based on chi-squared calculation, of the reliability levels and the density classes of discoveries (Fig. 10). The municipalities that received a very low reliability level are over-represented in the category in which municipalities have not yielded hoards or isolated finds and under-represented in the other categories. For the municipalities having a low reliability level, the values are near to what one can expect in case of independence. Lastly, the moderate and high reliability level are related to the municipalities having a relatively important density of discoveries (not necessary a very strong density, indeed these cases are the results of exceptional circumstances).

3.4 Creation of a “Representation Map” and a “Confidence Map”

A representation map aims at identifying the under- and over-representations in the data distribution (Oštir et al. 2008). The real number of discoveries is compared to the theoretical number in case of homogenous distribution. The following three steps were realized:

1 – Estimation of the expected number of discoveries per municipality after its area:

\[
(\text{Expected Nb. Disc.}) = (\text{Nb. Disc.}) / (\text{Tot. area}) \times (\text{Municip. area})^6.
\]

2 – Calculation of the difference between the real value and the expected value:

\[
(\text{Real Nb. Disc.}) – (\text{Expected Nb. Disc.}).
\]

3 – Classification of the values into 5 categories, corresponding to heavy under-representation (the data are absent or quasi absent), under-representation, normal

---

\[^{6}\text{Where: "Nb. Disc." signifies the number of discoveries, "Tot. Area" means the total area of the studied zone and "Municip. Area" stands for the area of the municipality in consideration respectively.}\]
representation, over-representation and exceptional representation (Fig. 11).

Finally, the confidence map is the combination of the representation and the reliability map (Fig. 12). It is the final tool that allows commenting the results of the analyses. This method enables more particularly to distinguish the documentary lacks for under-represented zones where there were not enough possibilities of discovery. This is a major element for understanding the results of the analyses. This permits in particular pointing out the zones about which no reliable archaeological conclusion can be made.

The reliability map highlights as well the zones with high potential that yielded a lot of discoveries. These zones are mainly situated along the Seine valley and the Oise valley. It can be supposed that they are well known since several factors could have been the occasion for the discovery of a hoard or an isolated metal find and they are well represented indeed. Nevertheless, does this high density correspond to a real archaeological tendency (more densely populated region or more frequent hoard deposition) or only to a high potential for discoveries that creates an over-representation of the data? That would be always difficult to make a distinction.

Finally, some zones can show an archaeological reality more recognizable and more reliable: namely where the impression given by the vestiges is different from the potential for discoveries. In the zones of low potential, while the archaeological discoveries are significant (large hoards, many isolated finds) or their quantity is important, it can be tempting to think that even a greater number of discoveries could be made there, but it still remains difficult to distinguish between the archaeological reality and the fortuity of discoveries. Lastly, the reliable zones that have yielded only few vestiges while their potential for discoveries is strong, credibly reflect archaeologically empty zones.

4. Conclusions

Generally, it is normal that the image given by the discoveries corresponds quite well to the potential for archaeological discoveries of each municipality. A discovery implies indeed that there was a possibility to find it. But what is more difficult to know and to estimate is how many times the digging up the soil did not yield any discoveries. In addition, the fortune’s role also should be taken into evidence. We can only catch its influence through the “exceptional” discoveries made in the zones of low potential.

It is obvious that discoveries do not have the same value regarding the potential of each zone. On a territory slightly dug, it is normal that we find only a few vestiges, so an important quantity of finds reveals a particular archaeological phenomenon. Conversely, on a territory heavily dug we normally find a relatively or even a very important quantity of finds. It is not easy to determine whether the important quantity of discoveries is linked uniquely to the richness in archaeological vestiges or if it is an effect of the over-representation created by a privileged data recording. Their statistical weakness reveals as well a particular archaeological phenomenon: a real archaeological void that should be taken into consideration as information about the space.

So we should readjust the value of the discoveries according to the potential of each
zone. The proposed confidence map has a modest ambition: it only intends to help to comment data distributions; the next step might be the difficult task of balancing the number of discoveries, according to the potential of each zone.

References


Assessing Positional Uncertainty due to Polygon-to-Point Collapse in the Cartographic Modelling of Archaeological Scatters

Fernando Sanchez and Antoni Canals
Universitat Rovira i Virgili, Spain

Abstract:
Positional uncertainty refers to ambiguity, inaccuracy and low precision with regard to the location of an event in space and/or time. Here we consider the positional uncertainty that arises with the polygon-to-point collapse operation in the context of data generalisation, and we use the spatial distribution of Middle Palaeolithic lithic surface scatters in the archaeological area of Sierra de Atapuerca (Burgos, Spain) as a case study. Sources of positional uncertainty are first referred to, before going on to present a technique to measure positional ambiguity based on the correlation between alternative collapse patterns.

Keywords: Polygon-to-Point Collapse, Positional Uncertainty, Density, Sierra de Atapuerca, Lithic Scatters, Landscape Archaeology

1. Introduction: the Polygon-to-Point Collapse Operation and Positional Issues in Geospatial Analysis

A commonly applied operation in cartographic generalization is the reduction of the number of geometric dimensions of an object, a data modelling solution known as collapse. Benefits from this transformation in spatial analysis are found both in the visual representation of phenomena and in drawing inferences on the spatial processes that determine their arrangement, as it allows the application of techniques based on relatively simple geometrical entities, typically point objects, to features with extent, which are usually modelled as polygons at the time of data capture (Illian et al. 2008, Longley et al. 2010). Such a degree of abstraction involves a major issue inasmuch as the spatial configuration of the original pattern is to be approached reliably, as there exists ambiguity as to where to locate the surrogate geometry. In this respect, three questions arise: (i) Should a particular object on a map not be collapsed, on the basis of the degree of positional ambiguity it generates? (ii) How much positional uncertainty is introduced by each collapsed object in the outcome pattern? (iii) How variable from each other are the alternative outcome patterns that result from applying different collapse rules to the same collection of objects? In short, if alternative collapse outcomes of the same original pattern resulted in spatial configurations that were different front each other beyond some ambiguity threshold, then the collapse operation might not be regarded as suitable.

Intuitively, the likelihood of realizing significantly different point configurations out of a single collection of polygons \( \{B_1, \ldots , B_n\} \) in a window of analysis \( W \) is connected (i) to the sizes of each \( B \) relative to that of \( W \), as the importance of the differences between alternative outcomes is determined by the scale of \( W \) (i.e. the same difference will be given a variable importance depending of the size of the study area), and (ii) by the shapes of each \( B \) relative to that of \( W \), as the spatial domain of the least compact forms could stretch across different sectors of \( W \). The relative size of \( B \) can be computed as the ratio of \( v(B) \) to \( v(W) \), where \( v \) denotes the Lebesgue measure of these two-dimensional entities (that is, their area size), whereas shape could be described by some compactness index (see below).

Corresponding author: fernando.sanchez@urv.cat
2. The Role of Point Density in Evaluating Configuration Ambiguity in Patterns of Polygon-to-Point Collapsed Objects

Exploratory spatial data analysis can be of help in seeking to solve the above mentioned uncertainties. The technique that here is proposed involves transforming the outcome of the collapse, which is a point pattern, into a real-valued model, a density field, with the goal of subjecting it to correlation measures. The procedure goes through three steps:

- Generate $n$ alternative point outcomes by applying $n$ different collapse rules to the initial set of polygons. In this regard note that the number of alternative point patterns is infinite as the number of points in a polygon so is, so it could prove more effective to sample the set of all possible alternatives and find a particular polygon-to-point collapse rule that minimizes the differences between the resultant pattern and the rest of alternatives, in order to define a reliable benchmark of the polygon pattern.

- Apply an intensity or density estimator to the alternative point outcomes, so point patterns are mapped to either intensity or density fields (Fig. 1).

- Describe the correlation between all pairs of alternative density fields.

We remind that point density is a transformation of the inhomogeneous intensity of a point pattern which estimates the value of the probability density function at a given location in space\(^2\), so the distribution patterns displayed by both measures are equivalent.

When no prior assumption on the distribution of the density is made, it is typically estimated by a function (the kernel) that takes information from the surroundings of the targeted location to give a density value, provided the sum of all density points in the data space equals to 1 (Silverman, 1986). Clearly, the larger the neighbourhood (alternatively, the longer the bandwidth of the kernel, in specialized parlance) the larger the amount of information that its read from the surroundings and incidentally the more similar the amount of point events that are taken into account from one subregion to another, so the greater the internal homogeneity or smoothing of the resulting density field. In other words, different bandwidths produce different density models from the same point pattern, so the alternative point outcomes should be subjected to several bandwidths in order to monitor how the correlation between fields varies as a function of kernel smoothing.

Correlation can be explored in a graphical way by bivariate scatter plotting, where a

---

\(^2\) Point intensity is the expected number of point events per unit area, which is equal to the total number of point events in $W$ divided by the size of $W$ in a stationary point process (Illian et al., 2008).
perfect match between two phenomena is depicted by points arranged along a diagonal ascending across the data space (Fig. 2). For many pairings of density field realizations this technique soon becomes impractical, however. In such a case we suggest to accumulate all scatter plots into the same data space so that variation among the correlation profiles of all pairs of alternatives can be explored jointly (Fig. 3).

3. Use of Reference Patterns for the Assessment of Positional Uncertainty in Collapsed Area Features

In order to control how correlation between alternative density fields behave under different degrees of positional ambiguity, variability of this characteristic has been monitored in a series of polygon sets with known morphometric features, which will later be used as reference polygon patterns. These consist of collections of non-overlapping, equilateral and equal-size triangular polygons that are regularly arranged within the unit area. Choosing this geometry as the reference simply responds to the fact that we seek to maximize compactness while minimizing the size of the reference geometry, in order to optimize uncertainty minimization on the basis of the two sources of positional ambiguity referred to above, the relative size and shape of the objects. It turns out that, out of the most compact shapes (such as the circle, the square or some cases of diamond-like rhombi), the equilateral triangle yields less ambiguity between alternative point outcomes if compared to polygons of a different shape but with the same perimeter length, as it combines a convex form with a relatively smaller area.
Figure 4 shows the relative size cases that are considered on designing the reference polygon structures. Next, for each relative size case the polygons are collapsed a number of times \((n = 100)\) so a sample of alternative point patterns is obtained, one of them being the mean centres of the triangles whereas the remaining ones solve the location of the centres by a random process constrained to the domain of the polygons, concretely a binomial point process (Illian et al. 2008). Since the mean centre tends to minimize the maximum difference between its location and those of randomly-generated centres, the point pattern formed by this collapsing rule ultimately tends to minimize its differences with the rest of alternatives patterns so it can be deemed a good surrogate of the universe of possible collapse outcomes.

The following steps involve transforming the alternative point outcomes into density fields and graphically investigating the correlation between pairs of fields, with one axis always being the derivative of the mean centres whereas the other will take account of the randomly-generated patterns. Figure 5 shows the correlation signals of the reference polygon patterns for several degrees of smoothing, which have been obtained by applying an isotropic Gaussian kernel function with four different bandwidths (Diggle, 1985). As expected, the polygon-to-point collapse operation results in point configurations that are less different from each other as the size of the triangular objects relative to the size of \(W\) decreases. With regard to these polygon structures, point configurations could start being regarded as near identical to each other when \(100 \nu(B) \nu(W)^{-1} \approx 0.20\) and virtually identical when \(100 \nu(B) \nu(W)^{-1} \approx 0.1\), so polygons could then be collapsed into points with a negligible degree of positional uncertainty.

Note that these similarity patterns between alternative point outcomes, and therefore the above relative size thresholds that relate to them, just refer to the presented reference polygon structures, which are composed of equilateral triangles. In order to measure to what extent the geometrical characteristics of objects in a collection depart from those of the reference geometry, shape measures can suitably be used. In this respect the Compactness Ratio (CR) measures how
larger or smaller an area object is in relation to
a circle with the same perimeter length as the
object. Because the circle is the most compact
gometry, the measure hence describes how
compact a figure is. Formally, let $a_i$ denote the
area of polygon $i$, $\partial_i$ its perimeter length and
c_i the area of a circle with perimeter equal to
$\partial_i$. Then compute $CR = (a_i / c_i)^{0.5}$. This ratio is
not affected by the size of the polygon and has
a value of 1 for a circular domain and a range
in the interval $[0,1)$ for all the other shapes (de
Smith et al. 2010). For an equilateral triangle
$CR \sim 0.778$.

The dissimilarity patterns obtained from
the reference structures are next compared
to a real case, as we apply the measures just
presented to surface scatters composed of
Middle Palaeolithic stone tools.

4. Evaluating Cartographic Collapse
in Practice: Surface Scatters in the
Sierra de Atapuerca (Burgos, Spain)

4.1 The Archaeological Record of the Sierra de
Atapuerca

The Sierra de Atapuerca (Burgos, Spain)
is a hill formation in the Northern Sub-Plateau
of the Iberian Peninsula and is a constituting
element of the Douro Basin (Fig. 6). Between
1999 and 2003 the area underwent intensive
prospection by the University of Burgos
Department of Historical Sciences and
Geography, with the goal of inventorying the
open-air archaeological record of the middle
Arlanzón (see Navazo, 2006, for details on the
data collection process).

Figure 7 shows those spots within the
study region where Middle Palaeolithic surface
scatters have been recovered. These have been
simplified in the form of polygons bounding the
assemblages, as the exact locations of individual
objects were not recorded in the prospection phase but just the limits of the concentrations.
The study region region ($134 \text{ km}^2$) is delimited
by the river Vena to the north, and the river
Arlanzón to the south, with the Sierra proper
at the centre of the area. This spatial domain
has been defined within the framework of
the doctoral research project of the first
author, which aims to explore properties in
the palaeogeographical substratum of the Sierra de Atapuerca and the middle Arlanzón associated with the spatial distribution of the archaeological scatters. No additional Middle Palaeolithic archaeological evidence has been documented in the surveyed region beyond the rivers, with the exception of two localities 10 km to the south of the Sierra.

The archaeological areas are, rather, a mixture of surface scatters that have later been excavated, surface scatters with minor test excavation and surface scatters with no additional excavation. Several common features among these assemblages have been observed, nevertheless (Navazo, 2006). First of all, they are exclusively composed of stone tools, which could be representing an original depositional pattern or the consequence of different preservation processes among the initially deposited materials. Secondly, the technical characteristics of these objects conform to the Clark’s Mode 3 of lithic reduction (Clark, 1967), which is equated to the Middle Palaeolithic in the European context. Thirdly, the objects are mostly made of flint, with a marginal presence of quartzite, quartz, sandstone and limestone. Flint-bearing outcrops are found on the top of the Sierra and to the east and west of this formation, but never beyond the oval demarcated by the rivers (Benito-Calvo, 2004), that is, beyond the study region. Quartzite and quartz pebbles are also found in the valleys of these rivers and of minor streams. No other flint source has been documented within 15 km from the Sierra and these external materials have not been identified in the archaeological record (García-Antón et al. 2003), a fact that makes the study region an exceptional area within the Arlanzón basin insofar as it provides a combination of palaeoenvironmental properties (a cave environment, a varied supply of raw materials and a network of alluvial plains) that is unique within the prospected region and its vicinity. Finally, the pieces have undergone minor taphonomic alteration, which mainly consists in weathering by frost and biological agents. Edges and arrises appear quite fresh in most of the cases, so post-depositional displacement is assumed not to have been important in relation to the spatial scale defined by the extent of the study region.

4.2 Distribution of Positional Uncertainty Among the Archaeological Surfaces

Before evaluating the positional ambiguity of the collapsed surface scatters as a whole, a first factor that is investigated is how the bounding polygons that approach the spatial domains of the surface scatters depart from the geometrical characteristics of the reference geometry (that is, the equilateral triangle), as it allows us to measure how well the archaeological surfaces match the geometrical benchmarks and hence how much ambiguity they contribute to the point pattern individually. To do so the relative size \(100v(B) / v(W)\) and the Compactness Ratio (CR) of the bounding polygons are compared to those of the reference geometries. These measures are shown in Table 1. The relative size of the bounding polygons concentrate well below 0.1 with the exception of six large areas (labelled as 5, 8, 10, 19, 24), so most of the positional ambiguity in the collapsed pattern is estimated to be due to the morphometry of these archaeological surfaces. Because of the quite skewed distribution of relative size values, robust central tendency measures should better be used in estimating the expected signal of the distribution of values. In this sense the Huber M-estimator of the mean is computed with two techniques (Huber, 1981): (i) by specifying an initial value for \(\mu\) equal to the median and (ii) by using the Median Absolute Deviance as a reference of the spread of the expected value. Further, outliers in the distribution of values are removed by Winsorizing the set at 1 and 1.5 standard deviations. In all cases the expected relative size is located in the interval [0.033, 0.046], quite below the relative size of the reference geometry that generates virtually identical collapsed patterns (which is \(\sim 0.10\), as Figures 5c and 5d have shown).

As explained above, the size of the
geometries alone does not suffice for comparing a given target pattern to the reference pattern, since the compactness of the geometries, which is a function of their shape, also defines the degree of dissimilarity between alternative point allocations. In relation to the open-air archaeological pattern of the Sierra de Atapuerca, the joint departure of the archaeological surfaces from the relative size and from the shape of the reference polygon structure is shown in Figure 8. Red points indicate the area and CR of the reference polygon structures with the four smallest relative sizes ($100v(B)v(W)^{-1} \approx 0.639$ and below). Dotted and dashed lines indicate the relative sizes of the reference polygon structures whose derivative point patterns were regarded, respectively, as near identical and virtually identical. If the targeted geometries perfectly matched the spatial domain of one of the reference polygons, then they should appear superimposed on the corresponding reference signal (i.e. on one of the red points). Agreeing with the results of the $M$-estimators, most of the relative sizes of the archaeological surfaces are even smaller than the relative size of the reference polygon that generates point outcomes with virtually equal configurations. Further, they tend to aggregate close to the CR of the reference geometry ($\approx 0.778$), so the location of their point abstractions are expected to be solved in quite local subsets of the study area. Only three archaeological surfaces notably depart from the reference CR towards elongated shapes, the first of which is, besides, the second largest surface in the polygon set (Table 1).

### Table 1. Relative size $100v(B)v(W)^{-1}$ and Compactness Ratio of the archaeological surfaces.

<table>
<thead>
<tr>
<th>Loc</th>
<th>Relative Area</th>
<th>Compactness</th>
<th>Loc</th>
<th>Relative Area</th>
<th>Compactness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0526</td>
<td>0.853</td>
<td>15</td>
<td>0.0612</td>
<td>0.836</td>
<td></td>
</tr>
<tr>
<td>0.0680</td>
<td>10</td>
<td>0.0480</td>
<td>0.832</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0456</td>
<td>0.744</td>
<td>17</td>
<td>0.0684</td>
<td>0.756</td>
<td></td>
</tr>
<tr>
<td>0.0676</td>
<td>0.764</td>
<td>15</td>
<td>0.0616</td>
<td>0.717</td>
<td></td>
</tr>
<tr>
<td>0.0698</td>
<td>0.707</td>
<td>19</td>
<td>0.0603</td>
<td>0.763</td>
<td></td>
</tr>
<tr>
<td>0.0093</td>
<td>0.886</td>
<td>20</td>
<td>0.0547</td>
<td>0.770</td>
<td></td>
</tr>
<tr>
<td>0.0003</td>
<td>0.886</td>
<td>21</td>
<td>0.0076</td>
<td>0.754</td>
<td></td>
</tr>
<tr>
<td>0.7382</td>
<td>0.645</td>
<td>22</td>
<td>0.0176</td>
<td>0.862</td>
<td></td>
</tr>
<tr>
<td>0.0093</td>
<td>0.886</td>
<td>23</td>
<td>0.0160</td>
<td>0.531</td>
<td></td>
</tr>
<tr>
<td>0.5648</td>
<td>0.718</td>
<td>24</td>
<td>0.2414</td>
<td>0.777</td>
<td></td>
</tr>
<tr>
<td>0.0498</td>
<td>0.829</td>
<td>25</td>
<td>0.0093</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td>0.0093</td>
<td>0.886</td>
<td>26</td>
<td>0.0097</td>
<td>0.786</td>
<td></td>
</tr>
<tr>
<td>0.5775</td>
<td>0.734</td>
<td>27</td>
<td>0.0118</td>
<td>0.651</td>
<td></td>
</tr>
</tbody>
</table>

### Figure 8. Departure of the archaeological surfaces from the reference polygon patterns: Relative Area Size – Compactness Index relationship. Details in text.

#### 4.3 Bivariate Association between Alternative Point Density Fields

In order to assess how alternative point configurations vary from each other in the archaeological pattern, correlation between pairs of alternative density fields is now computed. 99 alternative point configurations of the collapsed archaeological surfaces have been generated by obtaining the $x$ and $y$ coordinates of the collapse points by a random point process constrained to the spatial domain of the respective surface. An additional configuration defined by the mean centres of the polygons is also included in the sample of alternative point outcomes, as it will tentatively be considered the surrogate of this sample (Fig. 9).

From each alternative point pattern four density models with different degrees of smoothing are derived. The density functions are fixed-bandwidth isotropic Gaussian kernels (Diggle, 1985) with identical characteristics except for their bandwidths, which have been defined by setting the standard deviation of the function at 578 m, 868 m, 1,157 m and 1,447 m off the origin of the kernel. Such bandwidths approximate those used for estimating the
density fields of the collapsed reference polygon patterns, which have been 0.05 metric units (MU), 0.075 MU, 0.1 MU and 0.125 MU in the unit area, by making their values equal to $\beta v(W)^{0.5}$, where $\beta$ is 0.05, 0.075, 0.1 or 0.125 (depending on the degree of smoothing to replicate) and $v(W)^{0.5}$ the square root of the area of $W$. Since $v(W) = 134$ km$^2$, then $\sigma = \{578, 868, 1157, 1447\}$. Beyond $\sigma = 1447$ m the density fields become so homogeneous across the spatial domain of $W$ that similarity between point patterns cannot be described correctly.

Figure 10 shows the accumulated scatter plots of pairs of alternative density fields for the four bandwidths, the y and x coordinates being, respectively, the density signals of the mean centres of the archaeological surfaces and of the randomly-collapsed point outcomes. The plots show quite linear positive correlation patterns between alternative density fields, with correlation profiles lying somewhere between the ambiguity thresholds that were proposed for near and virtually identical alternative configurations in the reference polygon structures (Fig. 5), a result that agrees with the degree of ambiguity previously predicted on the basis of the morphometric features of the archaeological surfaces (Fig. 8).

5. The Degree of Positional Uncertainty in Collapsed Archaeological Areas: Recapitulation and Discussion

Three questions have been formulated in the assessment of positional uncertainty due to polygon-to-point collapse: (i) whether a particular polygon feature should not be collapsed into a point object on the basis of the degree of positional ambiguity it generates, degree that has been made be a function of the relative size and compactness of the original object, (ii) how much positional uncertainty is introduced by each collapsed object in the outcome pattern, and (iii) how variable from each other are the alternative outcome patterns that result from applying different collapse rules to the same collection of objects. In relation to the archaeological surfaces from the Sierra de Atapuerca, the compactness of the polygon abstractions as well as their relative sizes (compared to the size of the study area) are in almost three quarters of the cases within the interval of values that would generate virtually identical alternative point patterns according to the morphometric characteristics of the presented reference geometries (Figs...
5 and 8), so these could be collapsed with minor positional uncertainty. As for the third question, even though over a quarter of the archaeological spots have morphometric features that depart from reference geometries with low positional ambiguity, the accumulated scatter plots of pairs of alternative density fields depict a quite consistent linear relationship, something that is observed even with the shortest kernel bandwidth (Fig. 10). These results inform about negligible differences between alternative collapse point abstractions of the archaeological surfaces documented in the Sierra de Atapuerca. Inasmuch as such a correlation pattern describes alternative density fields that, globally, are deemed equivalent to those obtained from the mean centres of the polygon abstractions this observation, in addition, allows us to confirm the mean centres as a reliable surrogate of the sample of randomly-generated collapse point configurations.

6. Conclusions

Patterns of similarity between alternative outcomes from a polygon-to-point collapse problem are attributed to variations both in the shapes of the polygon features and in their size relative to that of the study region. This work has presented a protocol to monitor the effects of these morphometric characteristics based on the correlation between collapse outcomes, and advocates for the use of polygon models with specific morphometry and known positional uncertainty as reference structures. With regard to comparing density maps from alternative point outcomes, the accumulated scatter plot has proven suitable for analysing the frequency of particular association values in the sample of pairs of alternative point patterns, and hence a way to jointly describe positional uncertainty in this sample. In summary, this work shows the importance of taking the size and shape of area features into account in the assessment of sources of uncertainty when simplifying spatial objects in data generalization.

References


Theoretical Space-Time Modelling of the Diffusion of Raw Materials and Manufactured Objects

Estelle Gauthier  
UMR CNRS 6249, Chrono-Environnement, France

Olivier Weller  
UMR CNRS 8215, Trajectoires, France

Jessica Giraud  
UMR CNRS 7041, ArScAn, France

Robin Brigand  
UMR CNRS 6249, Chrono-Environnement and Alexandru Ioan Cuza University, Romania

in collaboration with: Pierre Pétrequin and Maréva Gabillot
UMR CNRS 6249, Chrono-Environnement and UMR CNRS 6298, ARTeHIS, France

Abstract:
Workgroup 3 of ArchaeDyn II programme focuses its study on the diffusion systems of ancient products. In order to be able to structure data in GIS in an appropriate way, we propose a general theoretical modelling integrating the different components of the diffusion systems, and identifying their interactions and the factors affecting the location of products and their transfers. Three dimensions are considered: the Time, the Space and the Function of places. A product’s pathway can be apprehended efficiently by distinguishing spatial entities as well as functional entities. This modelling highlights the fact that the approach through the simple notion of “site” is not sufficient to study the products’ diffusion without taking into account the nature and the role of the places. On the basis of this model, we propose a conceptual data model (created with the help of the HBDS method) that will finally lead to the creation of a “three-dimensional” geodatabase.

Keywords:
Theoretical Modelling, Diffusion Systems, Life-Cycle, Place, Time, Space, Function, Conceptual Data Model, Geodatabase

The Research carried out by Workgroup 3 of the ArchaeDyn II programme focuses on spatial dynamics with respect to the production, circulation and consumption of raw materials and manufactured products in ancient times (Gauthier et al. 2012). Our final objective is to use available datasets about various raw materials (salt resources, copper, etc.) or finished objects (jade axeheads, bronze palstaves, grinding stones, etc.), dating mainly from the Neolithic to the Iron Age, to test accepted models, which are all too often based on market economies (e.g. distance models), and ultimately to propose new models that incorporate the social status of the objects under consideration. This paper is devoted to the research for a method of data structuring in order to meet this long-term objective.
To analyse how the archaeological object survived until now after being produced, transformed, used and transported, even at several occasions during their existence, it is necessary to reconstruct their history. Two notions appear to be essential: space and time. This article proposes a way of data structuring that would, in theory, allow to integrate them into a GIS, taking into account the different stages the artefact has known and different places through which it passed (insofar as it is possible to know such information). This is above all a theoretical exercise. However, we neither intend to make a conceptual work nor to return to the various archaeological theories to discuss whether they are to be followed or not. It is a modelling work that aims to propose a method for recording and primary processing data. The components of the diffusion systems (Gauthier et al. in press), their organization, their interactions and the factors influencing the location of products or their transfers, will be listed and organized into the model.

The novelty in this work does not lie in the concepts it puts forward (defined mostly in the 1970'), but in our proposition of their organisation, through the creation of a model in which these terms find their connections, and in its adaptability. Indeed, the suggested model is designed to be a general one which takes account of the different possible configurations, so that it can subsequently be adapted to the various materials or products under study and so that specific diffusion models could be proposed.

1. Defining and Modelling Diffusion Processes

The object of our study is the archaeological artefact; it may be either a raw material or a manufactured product. This artefact gives evidences for its use or at least for an intentional or potential use (although not necessarily for straightforward technical or economic purposes). The whole of these objects and materials forms the circulating stock (e.g. Needham 1998).

Three dimensions have to be considered in order to apprehend the diffusion systems: the Time, the Space and the Function of places.

1.1 TIME Factor: the Product’s Life-Cycle

We do not consider here the time as the absolute time of the production, neither as a phase linked to typology of the object, nor as the dating of the site where the object was found, but as the history of the object, its “life-time”: so it is a relative time, specific to each artefact.

The product’s life-time is composed at least of one entry into the circuit (the first step is always those of creation/acquisition: exploitation, extraction, production, fabrication, transformation) and one exit from the circuit by being lost, abandoned, deposited (Bradley 2000, 37), destroyed; or by being reintroduced into the circuit in some new form by recycling (e.g. Schiffer 1972, 158; Needham 1998, 289, Fig. 1). This life-time may be apprehended as a cycle, and we therefore speak of life-cycle4.

4 We find this notion under different appellations in the bibliography, but they correspond to very similar concepts: ‘use-life’ (Schiffer 1972, 159; Tringham 1994, 174-177, including buildings as finished artefacts; Shott 1998, 313-314); or ‘life history’ (Tringham 1994, 188; Gosden-Marshall 1999, 169-170); or ‘product cycle’ (Brunet et al. 1992, 140); or ‘biographies of artefacts’ (Bradley 2000; Tringham 1994, 188, who uses it also for places and constructions).
Theoretical Space-Time Modelling of the Diffusion of Raw Materials and Manufactured Objects
Estelle Gauthier et al.

But between these two essential steps, the product may have had a variable pathway, going through one or more stages of production, consumption and diffusion. In this model no intermediate stage is mandatory and the pathways may be varied, the product even moving backwards and forwards among the different possible stages.

These three domains may indeed be defined as possible but not compulsory stages in the life-cycle of a product (Fig. 2):

- Production includes the stages of exploitation, supply and transformation (Needham 1998, 287, Fig. 1; Bradley 2000, 40-41).

- Diffusion concerns the physical circulation of products, their distribution but also the transmission of models. The idea of trade and transfer between sites and individuals is also included (Bradley & Edmonds 1993, 12-13).

- Consumption corresponds to the use of the product but also to some other manipulations representing a form of economic or social use, particularly deposition (thesaurization, hoarding) (Bradley 2000, 37).

During its life-cycle, the object is ascribed one or more functions that may change over time. It may undergo transformations as well. It is also ascribed one or more values (social and/or economic, etc., e.g. Hodder 1974, 1982; Pydyn 2000, 8-11; see also the recent multidisciplinary researches of the JADE programme, Pétrequin et al. 2012), which may fluctuate in time and space, depending on the individuals who possess them or on those who would like to acquire them; these values are therefore subjective. The object may have changed hands via distribution, trade, gift, inheritance, theft, appropriation, etc.; we shall speak generally of transfers (Brunet et al. 1992, 489). The series of such transfers may entail the product being moved in space over time, which we call diffusion (or migration as used in Pumain and Saint-Julien 2001, 157).

It should be noted that the studied products -archaeological discoveries- necessarily come from the final stage (abandonment, loss, burying, etc.). They have therefore exited the circuit, although they do reflect their involvement in the lives of past societies, at one of the stages distinguished above. Hence they may therefore be used to track all or a part of their life-cycle, knowing that they may not be representative of all of the products of the same kind in the circuit (quantity, location, etc.), and that it may be difficult to perceive certain stages. Recycling allows the object to begin a new cycle, but we cannot always easily recognize these objects (and they could have been many more than the number of objects discovered). C. Renfrew (1975, 41) warns in particular against the temptation of assuming a close correlation between the number of archaeological artefacts found and the intensity of use of the products.

1.2 SPACE Factor: Geographical Entities and Diffusion Networks

Although, from a thematic point of view, the object of the study is archaeological artefacts, the analysis of the diffusion processes will concern not the products themselves, but...
the places where they were found (where they exited the circuit) or places they might have transited (e.g. a mine or a workshop). These loci are usually referred as “sites” (“place” according to Binford 1982, 5-6).

The loci involved in the diffusion processes are thus spatial entities localized by coordinates in a defined system. On the scale of our study, they are generally spatialized in the form of point entities. All of the spatial entities through which the products pass successively form diffusion networks (Fig. 3). These networks may be materialized as graphs (Berge 1967, 70; Pumain and Saint-Julien 1997, 93), each vertex (node) being a place that is served, while the edges are polylines corresponding to the routes, lanes, tracks, roads and networks used for diffusion.

Diffusion corresponds to the gradual transmission and adoption of innovation over time and space, either by expansion, or by relocation (Bailly and Béguin 1998, 50). This definition may be applied to the diffusion of products (raw materials, manufactured items, models, types, symbols, etc.): it may be thought of as a set of transfers between different loci6. It implies a movement in space and therefore the passage of the products of a spatial entity to another.

In order to link the temporal model describing the object’s life-cycle and the spatial model underlining the nature of the diffusion networks, it is necessary to realise a third model that takes into account the function of the products and the places involved into the diffusion processes. This model, independent from the spatial locations, does not consider the loci as geographical but as functional entities.

1.3 FUNCTION Factor: Role and Nature of the Places Involved in the Products’ Life-Cycle

The different places (functional entities) potentially involved in the processes of diffusion in each of the three domains are (Fig. 4):

- Natural deposits (mines, quarries, etc.) or places of exploitation.
- Workshops or places of transformation.
- Intermediaries as places of distribution or redistribution (markets, places of transit, certain hoards): the product is supposed to stay briefly and be involved in a new transfer.
- Places of consumption, of use, but also of final deposition,8 loss or abandonment.

Each locus corresponds then to a function in the object’s life-cycle (‘functional entities’, which are generally named ‘contexts’), of which nature and role should be identified.

---

6 Here transfers within places are ignored if for identical contexts, e.g. between consumers in the same hamlet.

7 For example the ‘central places’ of C. Renfrew (1975, 5-6) would fit into this category for some of their functions. See also the idea of redistribution centre (Renfrew 1975, 10).

8 The act of putting artefacts deliberately and definitely in a closed space (e.g. a hoard or a grave).
Thus the same spatial entity having served for exploiting the raw material and the fabrication of the finished object, for its use and for its hoarding occupies different places in our model, and corresponds to several functional entities, for it concerns different instants of the product’s life-cycle.

The different stages of diffusion correspond to actions of different kinds (Fig. 5). Some of them are made within a unique functional entity (exploitation, production, transformation, transit, accumulation, exchange, utilisation, hoarding, thesaurization, abandonment, loss etc.). Others connect several functional entities: supply, put into circulation, transport, transfer, (re-)distribution, recycling, (re-)transformation...

1.4 Connection of the 3 Variables: Time-Function-Space

Considering the different types of functional entities involved in the 3 domains constituting the stages of the product’s life-cycle and the nature of the various actions that the products are subjected to, it is possible to combine temporal and functional models into a single model (Fig. 6). Each archaeological object can then be situated in the model, by knowing the stage at which it is in its life-cycle. Sometimes several stages through which it has passed can also be reconstituted.

Finally, if we reintegrate the notion of space, each spatial entity could correspond to different types of functional entities (Fig. 7). For a given object, each spatial entity can also represent successively different functional entities at each stage of the life-cycle of the product under study.

The passage from one spatial entity to another corresponds to a movement characterized by actions specific to the nature of the functional entities involved (Fig. 8).

---

9 However not necessarily different spatial entities; the notion of space is not considered in this part of the model.
Indeed, local actions do not entail change of spatial entity (and therefore no movement) even in case of change of functional entity. The same spatial entity may be the subject of a simple change of function in the product’s life-cycle (here the distance between functional entities = 0). But other actions imply movements: supply, distribution, redistribution, exchange, etc. Conversely, these transfers take the form of a circulation whenever the product moves through different spatial entities (greater than zero distance between the functional entities of source and destination). The transfers may be realized through certain place types without necessarily going through all of them in the proposed model; they may also be moved in succession through places of the same kind. The whole of such transfers are the flux (Pumain and Saint-Julien 2001, 183). In our case and for the time scale of our analysis, this may correspond either to the total amount of goods that have circulated or to the evolution over the different phases of the global distribution of quantities of product in a given region.

This modelling has highlighted the fact that the meaning of the usual term of ‘place’ isn’t sufficiently precise to characterize the phenomena under study. Indeed, it is required to distinguish loci as elements constituting the space (materialized by layers of point, linear or polygonal features) and loci as contexts in the products’ life-cycle. The approach through the simple notion of ‘site’ will not be appropriate to make a detailed study of the modes of products’ diffusion without taking into account the nature of these loci and of the actions that the objects were subjected to when they were produced, when they circulated and when they were consumed and without separating within these sites the different roles that they played in the product’s life-cycle. In order to apprehend efficiently a product’s pathway, it is important to distinguish, on the one hand, functional entities corresponding to the role played by the loci in the product’s life-cycle and, on the other hand, the spatial entities which are the locality and axes through which the product passed. In order to reconstitute the history of the archaeological artefact, the transfers between functional entities without necessary displacements would be taken into account as the “physical” pathway of the object during its migrations.

2. An Example of Application of this Modelling on a Theoretical Case

For a product under study, the first stage in the spatial analysis10 must be the identification and characterization of the different places having played a role in its life-cycle and the determination of the nature of the actions performed locally. The next one must identify the relations between the places involved in diffusion process (Pumain and Saint-Julien 2001, 8) by reconstructing the pathway followed by the product in the diffusion network and its chronology. The following stage furthers our understanding of the modes of diffusion by providing insight into the causes and mechanisms of diffusion. One issue would be to examine more closely the connection between social organization and transfer mechanisms.

Above all, the first step is the creation of the model for the objects under study. We propose in an absolutely theoretical example to present some of the many possible scenarios for the diffusion of a batch of products (Fig. 9). The model allows here at the same time

---

10 After creating a confidence map (Oštir et al., 2008; Gauthier et al. 2012, 66-67).
the reconstruction of the history and the pathway of the products. In this case, five point geographical entities are implied: two places of exploitation, one place of transformation, one of consumption and one of hoarding. Since the transfers of products are made between different places, there must be routes (linear entities) allowing their movement. The workshop is supplied with raw materials from both sources and it is to diffuse its production towards the places of consumption. The treatment of the raw materials makes some waste which can be abandoned on the spot. When the use of a product comes to an end either it is reintroduced into the circuit as a raw material or it can be placed in a grave or a hoard. Due to incidents during the transport, a part of the production can be lost on the way, e.g. in case of a shipwreck. In the end, the objects of this model (or at least there raw material) can enter and then exit the circuit after having passed through 2 to 6 stages (marked with “T” in the model) depending on the vagaries of its life-cycle. Of course a lot of other possibilities are not presented here, e.g., the consumption or abandonment at the place that served for its transformation.

3. From Model to Geodatabase

Our aim in the long term is the modelling of the spatial dynamics of diffusion processes (in their different aspects concerning the production, the consumption and the circulation) of specific archaeological objects. It would be ideal of course to work from data collected for this purpose, so that they are formatted properly to address these specific issues and to control their quality. But the establishment of such a dataset would require an important investment of several years. The challenge given to the Workgroup 3 of ArchaeDyn II was to work with pre-existing databases (13 in number, concerning different materials, different periods and regions, from a micro-region to the whole of Western Europe), already constituted in the framework of other research programmes. These are very heterogeneous data (geographical extent, scale, period concerned, exhaustivity of the inventory, reliability of the spatial repartition, etc.) and they are not formatted as we need.

The databases that we dispose are generally constituted in form of an inventory of artefacts (final products or raw materials) with a more or less precise spatial location (at best at the level of a given site, most often on the centre of the municipality that yielded the discovery). Others have the form of a corpus of archaeological sites having furnished data relating the studied questions and comprising some summarized information (number of objects, total weight, proportions of different materials, etc.). These data were recorded in simple tables. Attributes vary according to the product under study, and the primary objectives of the creators of the corpus (e.g. weight, level of finishing, traces of use, traces of voluntary breaking, context). Space is recorded only as an attribute of the sites or the objects (just the localization of the discovery), as well as the time which is indicated generally in form of a phase determined by the typo-chronology.

These two different manners in which the data are structured neither consider all the places implied nor the entire career of the object. These databases cannot be used, in their actual form, to reconstruct the history.
and the pathway of a particular object or the system of diffusion corresponding to all these courses. They are not directly exploitable for our purposes.

The model that we propose can serve as a basis for the creation of a conceptual data model which will then allow constructing a geodatabase specific to each dataset (Pirot and Saint-Gérard 2004; Saint-Gérard 2005). It will integrate the data in a standardised way, taking into account all stages of the life-cycle of the object and not just the final known stage. Space will no longer be considered as an attribute, the topology will be integrated into the internal structure of the spatial information, topological space supporting the data (Bouillé 1977; Pirot and Saint-Gérard 2004, 2005). Through the creation of a ‘three-dimensional’ geodatabase that takes into account both space (places), time (stages of life-cycle) and the function of the place (production, consumption...), we believe that we will achieve a satisfactory result in the long-term. We are only in a theoretical conceptual phase and we have not yet completed this geodatabase with our own data.

It’s necessary to decompose the information: this “three-dimensional” geodatabase may include each spatial entity corresponding to each functional entity in each stage of the life-cycle of the object in consideration. Thus, we should first reconstruct and model the biography of each artefact (Kopytoff 1986, 66-68) and then integrate it into a more complex model. But the reconstruction of this history for each object only based on archaeological evidences is a complex undertaking. The gaps are numerous. Even if it is possible to reconstruct some of the stages, generally those of the abandonment or of the production, we have a long way to go before we can reconstruct all the details of the ancient products’ diffusion system. The model should take into account the missing stages, places and actions, when they are known (but of course they will not be spatialized).

Concretely it is to translate the model we propose in the form of a ‘conceptual data model’, created using the HBDS method developed by F. Bouillé. This model is then translated into ‘logical data model’ and then into a ‘physical data model’, that is to say into a geodatabase implemented in ArcGIS (ESRI) (Pirot and Saint-Gérard 2004).

The proposed conceptual model (Fig. 10) is composed of two main parts: the first one corresponds to the spatio-functional entities while the second one includes the temporal and thematic entities (i.e. the raw materials or manufacture objects under study (or other traces) and the different stages of their life-cycle). The different semantic and/or spatial elements forming the context (not detailed here) depend on the product, period and region studied. The state of the research is taken into account as well to evaluate the quality of the data and to allow us to comment the spatial repartitions. Of course, the spatial entities and spatio-functional entities are linked by a one-to-many relationship and it is the same for the artefacts studied and stages of their life-cycle. The spatial entities bear general attributes of the sites; as the class of artefacts do so about the products. Each stage of the life-cycle will be spatialized via a spatio-functional entity (one-to-one relationship) in function of the role of this place in the life-history of the object and the nature of the action carried out on this object at this loci. So this class includes the attributes that allow establishing this link.

The geodatabase realized on this model varies depending on the theme of the study. Nevertheless, we propose here an example of its structure (Fig. 11). Hyperclasses concerning the context and relating to spatial data constitute different feature datasets (the examples given are neither obligatory nor exclusive). A specific feature dataset contains feature classes corresponding to reliability, representation and confidence maps. Another feature dataset corresponds to the different types of functional entities, while spatial entities form a separated
feature class (or may stay in a simple table). The corpus of artefacts is not itself spatialized, it is integrated in form of a table to which the table of the different stages of the life-cycle of each object is linked. It is through the link between the stages of the life-cycle and the functional entities that the spatial dimension of the stages of the products’ life-cycle is integrated. In fact, it is this link which is the most important and innovative in our approach.

This geodatabase will certainly prove to be an efficient way of structuring data. It is adaptable to each product, period and region. It remains, however, two questions about which internal reflection should be conducted. How should a place be considered: as a municipality, a site, a building? Therefore, what can be considered as intra or inter site? How can the problems of recording and accuracy of the data localization be resolved? These questions depend on the scale at which we intend to study the phenomena. In our case, given the scale of our study (from regional to European regional) and the average precision of the data localization (municipality names), the centroids of the municipalities will be considered as our points of reference. So sites and structures in the same municipality will be considered together in the same place. A second major problem persists despite our efforts to structure the data: the question of time. There is a fundamental difference between the scale of the dating of the object and the dating of the site. Our data (concerning Neolithic and Bronze Age) are recorded in phases lasting from 150 to 500 years (when it is actually possible to link the site to a given phase). We are therefore faced with a major problem of scale and precision of the dating sites and artefacts, which does not let us to locate in the absolute time the stages of the product’s life-cycle. It does not seem possible at this very moment of time to integrate it otherwise than as an attribute of the artefacts and the sites. This means it will not be, in the present state, possible to analyze precisely the spatio-temporal dynamics of products’ diffusion processes; the studies should be limited to the accuracy required by the precision of the dating of the sites.

References


Bouillé, F. 1977. *Un modèle universel de banque de


Theoretical Space-Time Modelling of the Diffusion of Raw Materials and Manufactured Objects
Estelle Gauthier et al.


A Tangible Chronology

Jean-Yves Blaise and Iwona Dudek
CNRS, France

Abstract:
A common task when trying to understand pieces of architecture inside a site is to spot, document and depict changes over time. Quite often successive states of an artefact are then represented using emergent, screen-based computer technologies (Virtual reality, Augmented reality, haptic interfaces, etc.). In this contribution we wish to investigate whether some tasks – both communication in workgroups or reasoning tasks – would not be better tackled once freed from the screen as unique interface. We introduce a proof-of-concept prototype, called “tangible chronology” developed in order to represent changes that occurred on Krakow’s market square over a period of 750 years. The paper presents the development and its evaluation, before discussing in what tangible models could serve content holders or academics specifically in historic sciences, and in what their making there calls specific attention and methods.

Keywords:
Architecture, Interfaces, Reasoning, Edutainment, Rapid Prototyping

1. Introduction

A common task when trying to understand pieces of architecture inside a site, whatever scale you choose to privilege, is to spot, document and depict their changes over time. This task pulls together different actors, with different agendas, and covers a rather heterogeneous set of challenges. For instance, from the point of view of scholars in historic sciences, the identification of a chronology means inputs, and inputs mean uncertainties, context, etc. On the contrary, content holders, magnetized by what they suppose are the expectations of the wide public; rather tend to minor doubts and to focus on communication tasks. But beyond these possible conflicts, there is a common issue: finding appropriate means to represent something that occurs, develops, and changes in time, and in space.

Ever since XIXth century pioneering works like Minard’s figurative cartography, or Marey’s graphic method, we are entitled to believe that depicting dynamics of change requires specific means and devices - in other words that the analysis of changes implies rethinking traditional visual tools like cartography, plans and section, etc. In the “information age” too, notably in visual analytics, innovative graphic solutions are put to the fore that renew our capacity to analyse and make decisions on spatio-temporal data sets (see for instance Keim et al. 2011).

However, when talking specifically about architecture, focus is still usually put on representing a state (“my building in 1605”), rather than changes. Readers probably came through historic edifices, sites, museums, exhibiting either coloured plans or virtual / tangible models showing various stages of development of the edifice. Naturally tangible models may now appear as “communication old timers”. And it is true that the development of computer applications (GIS, CAD, VR, the internet) and related technologies (LIDAR, laser scanning / 3D displays), have driven most actors to adapt their work methodologies. Seemingly, a time came (two decades ago?) when actors stopped thinking “what can you do with a virtual model that you can’t already do with a tangible plaster or wooden model”. Numerous investments in computer-based solutions later, it might appear rather absurd to consider there can be an alternative to screen-based communication.

Yet, our claim is that, with mature computer technologies and related devices
(touch screens, smartphones, etc.), with records of successes and failures, it might be time today to re-think this over calmly. In this contribution we intend to turn the question around: “what can we do with a tangible model that we can’t do with a virtual model”? In other words, we wish to demonstrate that maybe there are tasks – including reasoning tasks - that are best tackled once freed from the screen as unique interface. This paper therefore focuses on a clear research question: can physical models over-perform traditional screen-based virtual models (and naturally in which conditions, for which tasks)?

Our contribution introduces a proof-of-concept prototype, called “tangible chronology”, developed in order to represent changes that occurred on the market square of the city of Krakow (former capital of Poland) over a period of 750 years. The prototype combines a master board, 3D physical models of the artefacts, with a coding of their position on the board and in ordinal time, and a tangible timeline for each artefact (Fig. 1).

Naturally our aim is not to question the usefulness of computers in general: the tangible models we present come out of a fully computer-operated design and production process. Furthermore, tangible interfaces have become a hot research topic within the computer science discipline itself, and limits of the now traditional HCI devices (“mouse/keyboard/screen”) have long been discussed (and exceeded in the gaming industry). We shall therefore not discuss ergonomics on a general ground, but analyse in what tangible models can serve content holders or academics in historic sciences, and in what their making there calls specific attention and methods. In this paper we shall first comment on this global issue, present the prototype, and conclude with possible future directions.

2. Research Context

Over the past 20 years, we have seen the development of a number of technologies that have impacted the way we try to explicit architectural changes over time (changes inside an edifice, changes inside a site). (Mazuryk and Gervautz 1996) or (Novák-Marcinčin et al. 2009) give successive overviews of the some of the most commonly used technologies and CAD-based virtual environments such as VR (virtual reality) and AR (Augmented Reality). Although at start focused on design and analysis tasks, computer-based geometric modelling tools little by little integrated a concern for the visual quality of outputs, with realistic rendering layers introduced within many software packages. In short, what we have seen, notably in applications to historical sciences, is the focus progressively shifting from 3D for analysis to 3D for communication. This is particularly obvious when looking back today on contributions dating of the nineties like
(Alkhoven, 1993) : her analysis of the evolution of Heusden’s urban fabric, with a consistent classification effort, sheds light on geometric rules & knowledge in an analytical approach that has little to do with some communication-oriented approaches that the rise of technologies has made possible since then.

Under the influence of computer graphics – related fields, of mobile technologies, and of the gaming industry, new ways to deliver the same “virtual world” content have also been introduced – immersive platforms, haptic technology, touchscreen + stylus solutions, mobile devices, etc. It would be irrelevant here to draw the whole history of these technologies, but it appears to us important to state that seemingly, little by little, an idea has made its way inside many circles (wide public, content holders, funding agencies, some scientists, etc.) : if you want to illustrate how an artefact might have been at time t, how it might have been at time t+1, (etc.) , you need comprehensive 3D modelling, and virtual environments (the more the better). But what if you want to understand time, read & compare densities of changes, durations, overlapping, uncertainties? What if you’re interested not only in the effects of time on architectural shapes, but in how the whole story developed over time (Fig. 2).

As food for thinking and debate, let us take a classic example: virtual reality.

Is it that virtual? From the point of view of devices, gloves, helmets, captors are certainly not virtual. Now about the content – so-called virtual worlds. Are software, hardware, files virtual? Well in fact preserving such content on the long run is a recurrent issue, and not at all a virtual one. And so at the end of the day what is virtual in a virtual world? Well what is shown is virtual – pieces of architecture, reconstructed or not, i.e. virtual pieces of knowledge about an artefact’s evolution. Is this the goal of science?

And what has virtual reality to do with reality? If talking about realism in graphics, well then VR is as difficult to understand as reality itself. Furthermore, the actual reality analysts face when producing or processing historical data is partial indications, heterogeneity, uneven distribution of clues in time and space, etc. Is this the information virtual reality helps delivering? (Raposo and al 2008) comment this recurrent issue with this adroitly imprecise
quote from (Devine, 2007) “tension between authenticity and completeness” (sic.) Virtual reality might be “close to a reality”, but hardly to the reality of our knowledge on the evolution of artefacts. Our reality is packed with doubts, and therefore one should be cautious with what words like “VR technologies” covertly introduce in our practices.

Now we do acknowledge these arguments are a bit far-reached. But we believe technologies and corresponding practices are now mature enough to try and weigh cold-bloodedly their actual cost – no offence in that, I suppose. Our position is that 3D technologies in general – not only VR – not only at rendering time neither - invite us to say too much: Too much 3D, too many details, too many choices, too much rendering. And at the same time they do not allow us to say enough: not enough temporal aspects, not enough support for uncertainty and alternatives, not enough context. In other words these technologies, although convincing by many ways, might not be fully suited to the kind of pieces of knowledge and info we would like to communicate: they might not be least the only instrument we should use. And when indeed this instrument needs to be used, we then in return should better understand its cost in terms of relevance, readability and faithfulness in the context of historical analysis, and in terms of cognition for the end user.

3. Background and Related issues

Naturally, our intent is not to question the impact and potential benefits of emergent “virtual worlds” technologies, in particular since we also use them at times (Blaise and Dudek, 2009). Our intent is to discuss if for some purposes – workgroup discussions, abstract reasoning tasks, acceptability and usability for visually impaired people – we should not put more efforts on alternative solutions.

A full range of approaches do exist, in particular in the fields of infovis, positioned by (Friendly 2006), in terms of legacy, at the intersection of cartography, and statistics.

In cartography (Rød 2000) proposes a brilliant analysis of reality and its representation, quoting for instance Muercke: “the features on the map represent symbolic versions of reality, not reality itself”. And so a distance with reality would be welcome? (Delaunay 1995) says it in a clear-cut manner: the power of evocation of maps can be found in their capacity to reduce. To this day, at the scale of architecture, what this author calls a reduction process remains ill-formalised.

Moreover, when talking about historical sciences, time obviously plays a major role. But here again, at the scale of architecture, time plays a secondary role – with most often a series of states positioned along a timeline. The focus is usually put on “what did my artefact look like at this period of history, and then 200 years later” in a synchronic approach. Other aspects of the time parameter like densities, durations, and rhythms remain poorly dealt with. By contrast, key contributions have emerged in the visual analytics field on how to handle the parameter time in general terms (Aigner et al.2008), or in the specific context of uncertain data sets (Matousek et al.2007)(Blaise and Dudek, 2011).

4. Motivation and Challenges

Basically our motivation when this research started was to get closer to the reality we handle, and to find some media that, as an end-product, would be relatively free of unevenly mastered technological layers. Tangibility was for us only a track at start, fed by experiences showing physical models, provided that they are adapted to the public, can play a major role in communicating digital content in a more friendly way, as shown for instance by the InterANTARCTICA interactive museum installation presented by (de Bérigny Wall and Wang 2009) or in
(Thuvander et al.2008). And indeed, the recent development of rapid prototyping techniques, along with the adoption of de-facto geometrical standards stemming from the industry, opens unprecedented opportunities to rethink the role and impact of physical models. In addition, wireless communication technologies now enable easier interaction between physical models and digital content. It is also important to state that tangibility opens real opportunities for 3D content reuse – a challenge in itself (see Bilasco et al.2006) in particular when looking back on what remains of 20 years of geometric modelling in and around historic architecture.

Accordingly, hot research is today emerging on tangible interfaces for all (for the able-bodied as well as for the disabled people – see braillenet.org for instance) in cultural applications, but also in education (Scarlatos 2002), and more generally in interaction with multimedia content (Hirabashi et al. 2008).

And so at the end of the day tangibility appears as a promising solution, well suited for workgroup activities, relatively technology-free (as far as users are concerned), and what is more bringing in visually impaired people, and people with learning disabilities.

Yet in historical sciences time is key to understanding dynamics of change - and making time tangible is a challenge. Time to time transfers (transferring the 750 years evolution of an artefact into 750 seconds) are theoretically possible, but as shown by W.Aigner (Aigner et al.2008) ill-suited to analytical reasoning since no comparison is possible between events separated in time, or between large numbers of sequences. Furthermore, the readability of such a transfer for the wide public remains to be evaluated.

So to which extent is tangibility compatible with constraints one faces in historical sciences, namely reasoning on the parameter time, and handling uncertainty? A physical model can hardly represent the whole evolution of an artefact, and is by definition rather well-determined (at least in terms of geometry). This is precisely the issue behind the tangible chronology prototype: representing through physical models notions like changes (including non-morphological changes), durations and intervals, doubts. From this issue derive the experiment’s priorities:

- one product - several possible audiences,
- one product - several possible uses,
- communicate on temporal aspects,
- handle uncertainty in dating,
- support workgroup discussions,
- exhaustive spatial & temporal coverage on a test case.

The development and test of the prototype, as will be described in the following section, focuses on these specific challenges.

5. The Tangible Chronology Prototype

The prototype combines four elements: a masterboard on which 3D physical models of the various evolutions of edifices are positioned, chronocordes that act as a physical equivalent to timelines, and a carved variable values dice. Its specificity lies in tangible codifications used for various purposes, from positioning 3D physical models on the masterboard to confidence assessments. The initial concept, applied to Kraków’s medieval market square, was to offer users means to analyse visually and tangibly the site’s architectural composition, for any time slot of their choice inside the 750 years covered. To do so users combine on the masterboard 3D physical models corresponding to each edifice’s evolution at the period chosen. In other words, users are supposed to be given enough information to combine physically on the masterboard, and to analyse, all possible market squares between 1257 and today (this naturally implies working in discrete time, with a reasonable chronon - here one year - 754 combinations). Both spatial information (where was this edifice?) and temporal information
(did it exist at the date I am investigating?, if it did exist, which evolution did it reach?) have to be conveyed. Corresponding tangible codifications are at the heart of the prototype: in this section we first describe each code one by one, before presenting possible scenarios of use it finally comes in, and the test case.

5.1 Code 1: Positioning Edifices in Space on the Masterboard

The first thing needed is to position edifices in space, on the masterboard representing the market place. Each edifice is localised through a univocal and tangible geocode carved in positive beneath the physical model. The geocode is a simple 3 X 3 grid combining flat squares and thick cylinders, measuring approximately 1cm² allowing 2⁹ combinations. In addition to the grid, the geocode includes a small rectangle, 1cm long, that acts as a positioning pin at assembly time (in order to avoid mirror effects). For each edifice a corresponding geocode is carved in negative on the masterboard, allowing a univocal and fully tangible assembly (Fig. 3).

For almost each edifice, and therefore for almost each geocode, there are several

Figure 3. Concept and implementation of the geocode.

Figure 4. A number of physical models for each edifice, that need to be ordered in time.
3D physical models, corresponding to the successive morphological evolutions of the edifice (Fig. 4). There are for instance nine different 3D physical models for the town’s hall belfry – one of the three structures that remain standing up to now (the average number of physical model per edifice is between 3 and 4). And so we needed yet another code to distinguish the edifice’s first evolution from the second, the second from the third, etc.

5.2 Code 2: Dating Evolutions of Edifices in Ordered Time

Each evolution of each edifice is localised in ordered time (i.e. first, second, etc.) thanks to ordonators (cylindrical pins), a tangible codification carved in positive beneath the physical model. Corresponding slots are carved in negative on the masterboard, allowing users to check at a glance the overall number of known evolutions for each edifice (Fig. 5).

Taken together, geocode and ordonators are a univocal, time + space, patented assembly system. Yet evolutions of edifices are at this stage only ordered, not dated – no direct relation from edifice to edifice is possible, no mean is given to say which evolution was present at this or that year. A third codification was introduced to deliver more in-depth information on temporal aspects of each artefact’s lifeline.

5.3 Code 3: Localising Changes Inside Edifices in Discrete Time

For each edifice a sort of tangible timeline called chronocorde positions what happens and when in discrete time (chronon 1 year). A chronocorde features tangible codifications of dates, durations, key dates, uncertainty in dating, and differentiates functional transformations from morphological transformations. It is composed (Fig. 6) of shapes meaning time (square - turn of the century [a]; cylinder – decade [b]) and of shapes meaning events & processes (thick cylinders - functional transformations [c]; plates - dating of a morphological transformation [d]).
of events & processes, the position of the shape says “when”, the width of the shape says “for how long”.

Morphological transformations are dated by a time interval (2 dates in YYYY format). Plates representing these transformations combine three pieces of information (Fig. 7):

- The actual numerical dating (YYYY) is coded through small hemispherical shapes on both sides of the plate if the transformation lasted more than the one-year chronon, on one side only if it lasted less than one year.

- The top surface of the plates is engraved with gutter-shaped carvings so as to identify the order of appearance (ordered time) thereby allowing a direct relation of the plate to a given 3D physical model through ordonators (three ordonators <> three carvings).

- Finally, the foremost side surfaces of plates is carved with different iconic–like tangible shapes that represent a qualitative evaluation of the dating’s credibility (data stored as numerical scale).

5.4 Code 4: Learning to Say Less, Yet Say it All.

At this stage we now can position edifices in space, order then in time, date their changes with, to some extent, an uncertainty assessment. Yet, when talking about uncertainty, besides temporal aspects the morphology has to be taken into consideration (with unknowns of its own). In some cases, the information we have can help state “what kind of edifice” it was in terms of structure and usage but not exactly how many windows it had, or how steep the roof was. And so if we are to question the relevance of tangibility as a media, in the context of weak data sets, we need to learn saying not too much about edifices, with tangible means. Still, as a result of our first evaluation campaign, it appeared clearly that in the case of visually impaired people infos like “how many storeys”, “on which side was the entrance” or “what material” could be very useful to deliver in order to foster comparisons. Accordingly, we introduced codifications that say, broadly speaking, “all we know is...” and thereby help us get closer to saying all we know, and nothing more (Fig. 8). This clearly calls for further development – and is a challenging issue not only in the context of tangible models.

5.5 Scenarios of Use

Although some issues it raises do call the researcher’s attention, the tangible chronology prototype is primarily targeted at use for the wide public, in the context of museums and/or education. It was initialled developed to serve
in the context of expert-guided pedagogical activities planned by the MHK (historical museum of the city of Krakow) as part of its extension.

Elements presented here above can be combined in three scenarios of use: guided interactive workshops, standalone presentation, board game.

In the guided interactive workshops (museum and / or education) mode, users are invited to recompose the market square for any time since 1257 under the guidance of an expert who intervenes to comment on the context of the period, the reasons for changes, the historical sources available, the doubts that remain (Fig. 9). In this mode all the 3D physical models are laid out on an assembly table, along with their chronocordes. As will be discussed in the evaluation section, the whole installation is part of the educational benefit of the prototype. It favours what in infous is known as context + focus approach (giving an interactive access, in the same space, to an overall view the chronology of all edifices and to a detailed view on each evolution of each edifice). It has proven to perform extremely well in terms of support for workgroup discussion – an aspect we had absolutely not anticipated, by the way. Furthermore, unlike one-user computer-centred installations, it clearly encourages collaborative group interactions, and is suited to groups of various sizes. But naturally there is a cost to this benefit: such a use requires guidance. Accordingly, there might be more perspectives of applications in the educational field, as part of the pedagogical material through which the teaching is done, than in traditional museums.

The prototype can be also laid out as a standalone presentation: 3D models are then positioned with regards to the morphological
transformation plates of the chronocorde (Fig. 10). Focus is here put on comparative reading, combining temporal + morphological aspects edifice by edifice.

Tangible chronology is clearly an edutainment initiative, allowing users not only to perform data mining tasks by touch and vision, but also to perform intuitive assembling tasks (category 3 of the ESAR classification – Exercise play, Symbolic play, Assembly, games with Rules, see Garon et al. 2002) that are particularly fruitful for people that are emotionally vulnerable, or have learning disabilities. Accordingly, we tried to evaluate where the idea of playing with the prototype could lead. A board game was designed for which a dice with variable values was created. The baseline of the game’s scenario is to have gamers get hold of edifices, in relation with dice-dependent time slots, and build an “impossible city” (putting together things that did not coexist in time). This application however goes a bit beyond the scope of this paper and we therefore do not detail it further.

5.6 Test Case and Production Process

As mentioned, our experiment was carried out on the development of Krakow’s market square over the past 750 years. Up to 40 structures coexisted at one moment in time or another during that period inside the 200m x 200m market square – yet only three of them remain up to now. Moreover, most of these structures are known to have changed repeatedly, may it be changes in terms of morphology or in terms of usage.

The documentation, the morphology of each evolution, as well as the chronology resulted from comprehensive architectural analyses we carried out and published in the past years, with ad-hoc contributions from our colleagues W.Komorowski and T.Weclawowicz. This input comprised dynamic VRML models, formerly used as interfaces in an online documentation platform (Blaise and Dudek, 2005), and a structured description of changes that occurred on each and every edifice during their lifeline (Dudek and Blaise, 2010). Naturally levels of knowledge, and levels of confidence, strongly vary from edifice to edifice, both in terms of morphology and in terms of dating. Furthermore, inside one edifice’s chronology, doubts exist here and there, and by the way not necessarily for the oldest period, as demonstrated in (Blaise and Dudek, 2011). We transferred the VRML models into physical shapes with some additional CAD modelling needed to integrate the prototype’s codification (geocode, ordonators).

The actual physical models produced cover 25 structures, and 90 evolutions: we chose in a first stage to leave aside structures that are too weakly documented – their very existence as independent edifices being even questioned. Accordingly we do not claim we managed to transfer the whole indications we have gathered throughout the years to physical models – we did get closer to the reality of the knowledge we handle, but not exactly to it.

The prototype was produced using a rapid prototyping 3D printer from ZCorp. The codification was integrated inside the initial
dynamic VRML 3D models so as to avoid dependence to the 3D printer technology. Models were translated into STL format at production time in order to allow geometric checking in the printer-specific software (consistency checking, accuracy, etc.). The chronocorde element however remains “hand-modelled” at this stage – a transitory and unsatisfactory situation.

6. Evaluation

At mid-development we carried out a first feasibility / readability check with visually impaired people, blind people, people with learning disabilities, thanks to the cooperation of an association specialised on these issues in Kraków. The idea was to evaluate the whole concept and the tangible codifications through informal discussions. A number of positive points were made on the acceptability and accessibility of the prototype, which apparently was a real help for testers to verbalize questions and concern.

However the readability of the architecture itself was questioned at two levels: the size of the edifices “in the real world” (i.e. readability of scale), and a frustration that “you can’t touch what’s inside the edifice”. To the first concern we did react by proposing additional codifications (see section 5.4), but the second obviously requires a change of scale – something we are currently working on, and testing. More generally speaking, this first evaluation brought us to re-think the way we should try and make space sensible for blind people (typically for instance, direct geometric scaling and transfer from reality to 3D physical models can be totally inappropriate since they imply an a-priori knowledge of architectural spaces not necessarily part of their personal experience).

In a second round, we carried out a more structured evaluation with various age groups in a High school in Kraków (VIII PLO). Testers from three classes (ages 14 to 17) were first presented the various components of the prototype (masterboard, 3D physical models, chronocordes) and then left, unguided, with precise questions on the development of the market square. Questions required a clear understanding of the information conveyed by the above components. A typical exercise of the evaluation was for instance, given a painting, to date it by its architectural content using the tangible chronology prototype’s components (Fig. 11).

Finally, written comments on the overall system were compiled and discussed informally. Lessons from this evaluation can be summed up as follows:

• the prototype clearly favours questioning and interactivity from the audience,
• at this stage an initial guidance is needed,

• the prototype helps shedding light on the process (interpretation of historical clues) rather than on the results (some set of 3D shapes),

• it helps going backwards, back to hints, clues, methods, have them understood by the audience, raises the audience’s awareness of historical sciences,

• it uncovers differences in levels of knowledge between various elements in a data set, in our experiment an urban ensemble,

• It puts temporal aspects on equal terms with spatial aspects.

Additionally, and surprisingly in a way at a time when computer gaming is so strong, the board game use raised high interest. More generally, the evaluation proved that saying less, and in a tangible manner, through objects, means leaving more space for interaction and questioning, means more understanding, even fragmental, of what it looks like to face historical data sets – and therefore in what it is challenging, interesting.

7. Conclusions

This contribution was not about reaching to general conclusions on communicating science, but simply about reporting on an experiment through which we have tried to make the evolution of a site tangible. The tangible chronology prototype, initially designed as an edutainment platform in the context of museum activities, does however underline some more general issues:

Scientists from the Humanities fields are not condemned to be stamp givers for tech developers blinked by computation capacities and graphic realism.

In some cases, and in particular in workgroup discussions, traditional “3D Modelling + Virtual environment” solutions are over-verbose solutions, which besides inadequacy to our data prevent users from an appropriation of the challenges behind historical sciences.

In short, what the tangible chronology experiment and its evaluation show is that consistency with the underlying data’s specificity (temporal aspects, uncertainty) does not undermine scientific communication, it helps it.

It has to be said at this stage that for most of it we could have reached this conclusion without experimenting physical models. The prototype shows a succession of discrete states, and emphasizes time / uncertainty parameters – and this can naturally also be done through visual means. It is not said though that the demonstration that there is more to Historic Sciences than still life 3D m sy it be (interactive or, animated), probably would not have been as radical. In addition, the experiment has not yet come to its end. In a next step we need to integrate technologies for object-computer interaction (i.e. 3D physical models / chronocordes as tangible interfaces) in order to deliver multimedia digital content. But for this step to be taken (a step at hand in terms of technology) freeing oneself from the screen’s attraction is required from many...

References


Reconstructing Fragments: Shape Grammars and Archaeological Research

Myrsini Mamoli
Georgia Institute of Technology and Massachusetts Institute of Technology, USA

Terry Knight
Massachusetts Institute of Technology, USA

Abstract:
Reconstruction of fragmentary archaeological remains includes a high degree of uncertainty. Digital models are increasingly used to describe and test hypotheses about the reconstruction of these remains. Among them parametric and generative models are the most promising ones: Parametric models allow for variation based on an underlying structure and generative models allow for variation based on rules. Shape grammars and parametric shape grammars are suggested here as the best approach of generative models because they are based on visual rules allowing for ambiguity, multiple readings, and variant reconstructions and interpretations. The case study of classical libraries is used to exemplify how shape grammars can be used in archaeological analysis and reconstruction.

Keywords:
Shape grammars, Archaeological Reconstruction, Ancient Libraries, Generative Modelling

1. Introduction

Archaeological research increasingly relies on three-dimensional representational techniques including boundary modelling, point-cloud modelling, solid modelling, volume modelling, and other similar geometric methods. These modelling techniques are primarily used as descriptive tools to represent artifacts and/or stratigraphic material and only secondarily as analytical and evaluative tools to explore hypothetical reconstructions and alternative versions of the found artifacts. More recently, parametric modelling has been extensively used to describe and interpret artifacts, classes of artifacts, and stratigraphic material too. In this representational mode, the geometry of the found artifact is encoded in parameters in predefined schemas and evaluated vis-à-vis the sets of designs defined by the schema. Despite the relative success of parametric modelling in the effective description, interpretation and evaluation of classes of found artifacts, the limitations of the approach are easily seen: several artifacts may be found that do not fit predefined schemas, and several hypothetical representations may never be modeled because they cannot be modeled by the given schemas.

A nice way out of this conundrum is offered by another class of modelling methods and tools, the generative grammars and generative modelling at large, where the designs are constructed by combinations of parts following formal rules. In generative modelling, symbolic rules are defined and apply recursively to generate designs. The benefit of the grammars is that the underlying structure of the given artifact emerges through the computation and there is no need to be predefined in a schema. The problem with generative modelling and grammars is that it requires an extensive computational knowledge to encode rules in a software environment - a knowledge that archaeologists do not have, and even more, that when rules are represented as symbolic structures cannot capture emergent aspects of computation with shapes. Here we focus on a new methodology for the analysis and reconstruction of the archaeological material, based on shape grammars (Knight 1994; Stiny 2006, 1980, 1976; Chippindale 1992) and we
suggest new trajectories of their applications in archaeological discourse. Shape grammars are formal systems that use shape rules to apply recursively to generate designs. Using shapes rather than verbal or symbolic descriptions, shape grammars can reconstruct classes of artifacts and instances of classes that follow the same design principles. A form of shape grammars that generate small and large-scale architectural artifacts by subdivision have already been used in archaeology (Dylla et al. 2009; Maïm et al. 2007; Müller et al. 2006). A brief description of few applications of set (symbolic) grammars and shape grammars in archaeological research is given below and the emphasis is given in the critical evaluation of these formalisms in archaeological discourse.

2. Grammars in Archaeology

Grammars have been used in archaeology as a way of classification and ordering the mass of findings, and as a systematic tool of reconstruction that allows possible reconstructions and interpretations (Hodder and Hutson 2003). As archaeological research and excavation progresses and new material evidence comes to light, the grammars can be modified to account for the new designs, or the new material can verify hypothesized designs generated by the grammars.

The earliest examples of applications of grammars in archaeological research are symbolic grammars (Hodder 1982). Symbolic grammars consist of a vocabulary of symbols, a set of rules according to which the vocabulary of symbols is manipulated and an initial symbol that defines the starting point of the computation. The computation is based on recursion. The basic elements are sets of symbols, typically numbers or letters.

An early example of a symbolic grammar in archaeological discourse is the grammar of the decorative arts in the area of Nuba hills in central Sudan (Hodder 1982) that appear in the decoration of hut fronts, pots, calabash artifacts and body-painting. The grammar encodes constructively the type definition of Nuba designs by decomposing the underlying motif, the star motif (Fig. 1a), into a vocabulary of shapes (Fig. 1b), and their spatial relationships, and came up with a small set of commands (rotate, duplicate, etc) that can compute known and hypothetical designs (Fig. 1c). By knowing the actual among the possible designs, he was able to identify which underlying rules are tribe-specific, and which rules are common to all tribes of the Nuba area. Hodder’s conclusion was that if the grammar were programmed in a computer, it would be able to produce all possible designs, known and hypothetical alike.

Shape grammars (Knight 1994; Stiny 2006; 1980; 1976) is a formal system that works in a similar way to symbolic grammars but with the additional advantage that all data are spatial. More specifically, shape grammars consist of a vocabulary of shapes, a set of rules according to which the vocabulary of shapes is manipulated and an initial shape that defines the starting point of the computation. The computation is based on recursion and embedding. The basic elements are a set of points, lines, planes and solids defined by their boundaries, and not predefined primitives. The rules are of the form A -> B and are applied recursively. Every time the left side of a rule A is embedded under any transformation in part or all of the design C so that t(A) ≤ C , the part of C is substituted by the right part of the rule B as described by the formula [C - t(A)] + t(B). Every time a rule is applied, all shapes fuse and
the design is re-described as a set of maximal basic elements. Rules can work in an additive, divisional or subtractive way, parametrically or not. Designs generated by the same grammar share the characteristics and features built in the grammar. Clearly the fact that the rules are encoded in shapes and the computation is visual as opposed to a series of commands written in some script or programming language method is more intuitive to the archaeologist, who is able to evaluate its validity or identify discrepancies more easily than with symbolic and verbal descriptions.

An early example of a shape grammar in archaeological discourse is the grammar of megalithic tombs in the Orkney Islands (Chippindale 1992). The grammar analyzes and describes constructively the design and the process of generation of megalithic, chambered tombs constructed out of flat stone slabs in a linear arrangement. With six rules that add extra modules, add vertical slabs and terminate the design, the grammar describes the form of known tombs and the process of their construction in real life, something that had not been possible with traditional research.

Shape grammars are a useful tool in the formal analysis of style and they can be easily deployed for the formal analysis of transformation of style (Knight 1994). Within this framework, stylistic change can be nicely captured by transformational grammars that transform one to another by changes in the vocabularies of shapes, the spatial relations between these shapes, the transformations under which the rules apply, and the rules themselves respectively. A nice example of a shape grammar that describes stylistic change in archaeological discourse is the grammar for the Greek meander motif in geometric pottery (Weissman-Knight 1986). The grammar is designed in three stages to capture stylistic change due to temporal and geographic criteria.

In the first stage, the grammar consists of the initial shape and two rules (Figs 1a and 3b), that generate simple meanders of the early and middle geometric pottery (Fig. 3c). In a second stage, the grammar is transformed with rule addition to account for the more complex meanders of the late geometric pottery; A stacking rule allows the stacking of meanders and a rule deletes the overlapping lines to generate double and triple meanders.
Additionally, rule change with shifts of 0, \(\frac{1}{4}\) or \(\frac{1}{2}\) units account for meanders of different areas (Fig. 3d). In the third stage, the grammar is transformed to account for stylistic change based on the style of different workshops of the same area and same period.

3. A Case Study for Classical Libraries

It has been argued that shape grammars is a useful formalism for the description, interpretation and evaluation of artifacts in archaeological discourse and for capturing stylistic change too. Two trajectories of new applications of the formalism are readily suggested: a) a critical juxtaposition of the generative aspect of the formalism vs. the notion of the schema as an underlying configuration used to produce constructive instances of a style; and b) the usage of the formalism to derive complete representations out of parts – in other words, to help produce reconstruction models based on the partial evidence produced at the excavations. Both trajectories could use any example of the literature to take on. The case study here is the architectural form of classical libraries, Greek and Roman. The first trajectory will be briefly touched here using a comparative discussion of proposed schemas of libraries and their recast in a grammar of libraries. The second trajectory, a by-far more ambitious trajectory requiring exhaustive documentation and analysis of the given corpus is currently under development.

Among civic buildings in antiquity, libraries are the least easily identified and the history of their architectural form is still unresolved (Callmer 1944; Makowiecka 1978; Strocka 1981; Wendel 1949). The history of the architectural form of the ancient library is typically confined to a very small corpus

<table>
<thead>
<tr>
<th>Library</th>
<th>Location</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Library of Pergamon</td>
<td>Pergamon</td>
<td>200 – 175 BCE</td>
</tr>
<tr>
<td>Library at the Serapeion</td>
<td>Alexandria</td>
<td>300 – 250 BCE</td>
</tr>
<tr>
<td>Academy of Plato</td>
<td>Athens</td>
<td>Hellenistic times</td>
</tr>
<tr>
<td>Library of Rhodes</td>
<td>Rhodes</td>
<td>Hellenistic times</td>
</tr>
<tr>
<td>Augustan Palatine Library</td>
<td>Rome</td>
<td>28 BCE</td>
</tr>
<tr>
<td>Library at the Portico of Octavia</td>
<td>Rome</td>
<td>23 BCE</td>
</tr>
<tr>
<td>Library at the Temple of Peace</td>
<td>Rome</td>
<td>75 CE</td>
</tr>
<tr>
<td>Domitian’s Palatine Library</td>
<td>Rome</td>
<td>80 CE</td>
</tr>
<tr>
<td>Pantainos Library</td>
<td>Athens</td>
<td>102 CE</td>
</tr>
<tr>
<td>Ulpian Library</td>
<td>Rome</td>
<td>114 CE</td>
</tr>
<tr>
<td>Hadrian’s Library</td>
<td>Athens</td>
<td>131 CE</td>
</tr>
<tr>
<td>Celsus Library</td>
<td>Ephesus</td>
<td>After 117 CE</td>
</tr>
<tr>
<td>Neum Library</td>
<td>Sagalassos</td>
<td>After 120 CE</td>
</tr>
<tr>
<td>Melitine Library</td>
<td>Pergamon</td>
<td>After 125 CE</td>
</tr>
<tr>
<td>Library of Nyx</td>
<td>Nyx</td>
<td>2nd century CE</td>
</tr>
<tr>
<td>Library at the Forum of Philippi</td>
<td>Philippi</td>
<td>2nd century CE</td>
</tr>
<tr>
<td>Aquatius Library</td>
<td>Tingad</td>
<td>2nd half 2nd century CE</td>
</tr>
</tbody>
</table>

Table 1. Corpus of known and identified libraries.
of ancient writings and an equally small corpus of surviving monuments. There are only seventeen buildings that are known from written sources and have also been identified with building remains. These seventeen libraries constitute the corpus of buildings to be used in this work.

The corpus of libraries showcases a big variation in form, layout, and scale. Libraries were built either as part of complexes, or as complexes or independent buildings, and consisted of several spaces to accommodate their diverse functions, the storage and retrieval of text: the main hall, and secondary spaces for additional stacks, offices, recitation spaces, auditoria. A very important element of their design was the association with a courtyard, an exterior leisure space, surrounded by stoas. In Roman times, the design of the library was formalized and monumentalized with the introduction of a monumental interior design in the main Hall, with rectangular niches in the walls, more typically in two levels, where the books were stored in wooden cabinets inserted in them, and a podium, which supported a colonnade and a gallery or an entablature. In the center there was a focal point formed by an aedicula, an apse or an enlarged semicircular or rectangular niche. Fig. 4 gives the reconstruction plan of the known and identified libraries, as they have been published by their excavators. The building remains suggest that in some cases more than one reconstruction is possible.

**Figure 4.** Corpus of known and identified libraries: a) Library at the Serapeum; b) Library of Pergamon; c) Library of Rhodes; d) Academy of Plato; e) Augustan Palatine Library; f) Library at the Portico of Octavia; g) Library at the Temple of Peace; h) Domitian’s Palatine Library; i) Pantainos Library; j) Celsus Library; k) Neon Library; l) Library of Nysa; m) Melitine Library; n) Ulpian Library; o) Hadrian’s Library in Athens; p) Library at the Forum of Philippi; q) Rogatimus Library.
Archaeological research has tried to identify the evolution in the architectural form of the library and classify libraries chronologically and geographically (Callmer 1944; Makowiecka 1978; Strocka 1981; Wendel 1949). These approaches have identified the common underlying design principles between libraries, and they emphasized the differences between others, without explaining how one type appeared in different regions and different eras. Open research questions include questions about their classification, their origins, typological and functional relationship to other building types, stylistic characteristics, and issues of layout.

The approach of this work is to describe the variation in the architectural form of libraries formally, by showing the transformations under which the design principles of libraries apply in each case. A first attempt to formalize the architectural form of libraries was made by Makowiecka (Makowiecka 1978), who established the different types of Libraries into eight schemata (Fig. 5), to describe known and hypothetical libraries. However, Macowiecka’s work did not include all possible variant libraries. Her schemata give the false impression that characteristics included in one schema cannot appear in the others, therefore excluding combinations between schemata. For example, the Library of Rogatus in Timgad is a combination of the schema with the two storage rooms and the central semicircular space and the schema with the portico.

This work is not much different than this, but instead of summarizing the design principles of classical libraries with parametric schemata, it summarizes them with rule schemata defined by parameters and conditions, that avoid repetition, and still give the freedom of computing every possible variation of possible libraries – existing and hypothetical.

The library grammar is a two-dimensional shape grammar that works in plan for the general layout and the main hall of the library and is work in progress. The library grammar has as initial shape the outline of one hall only, the main hall of the library. The rule schemata are developed in two stages; in the first stage, rule schemata generate the general layout of a library, with the optional addition of extra rooms, exedras, stoas and courtyard, and in the second stage the interior of the main hall with the specific architectural features, the podium the niches, the focal point and the entrance.

A short outline of the grammar that generates the libraries that follow the design principles as exemplified in Makowiecka’s schemata is given here to demonstrate how a generative grammar is a better approach than parametric schemata (Fig. 6). Specific parameters, conditions and labels in rules are omitted here for clarity. The first five rule schemata add additional rules, and/or duplicate the main hall and add a peristyle with a courtyard. The rule schemata 6-18 generate the interior of the main hall; modify the back wall or the whole main hall into a

![Figure 5. Schemata of Roman Libraries (Makowiecka 1978, redrawn after fig. 18-25).](image-url)
In the full version of the grammar, more details are given and the main hall is generated in plan and elevation to account for the variation in the podium, the niches, and the focal point. Also, more alternative rules account for more diverse layouts of different topologies and scales; single room buildings, buildings with liner, L-shape, or U-shaped stoas, and peristyles and exedras and other rooms added to them.

The shape grammar is parametric and is based on the analysis of the corpus of known and identified libraries, as it has been documented in a database. The rule schemata, their parameters and the range of valid values in the parameters are all derived from the dataset of known cases. Attached to the rule schemata are metadata that give the number of occurrences of each rule in the archaeological record. As research progresses and more evidence comes to light, the grammar can be modified to account for it.

The shape grammar generates a large set of designs, and works as a tool for reconstruction, evaluation and prediction. By starting from the building remains of an identified and known library, and by embedding the rules in its fragmentary building remains, one can reach variant possible reconstructions for each library. In addition, one can follow the same process with any building remains and evaluate their interpretation as a library. If it is possible

Figure 6. The library Grammar that generates the general layout and the main hall of the libraries presented in the schemas of Makowiecka and other hypothetical ones. Rules 1-5 generate the layout of libraries, Rules 6-8 generate the plan of the main hall, Rules 9-12 generate niches, and Rules 13-18 generate the openings to the main hall.

Figure 7. Derivation of a library as given in the schematas of Makowiecka.

Figure 8. Derivation of a possible library not included in the parametric schemata of Makowiecka.
to embed the rules in the given building remains and produce a reconstruction, then the building remains can be identified with a library. Last, one can start from scratch, from the initial shape of the grammar and by applying the rules to generate a large set of hypothetical designs that follow the same design principles and that future archaeological excavation could bring to light one day.

4. Discussion

Shape grammars, a formal method of visual computation, is appropriate for archaeology in the reconstruction of fragmentary architectural remains. It is a method that forces the researcher to explicitly articulate statements about the parts and the principles in their configuration, and encode them in visual rules that can be tested against the data. The fact that the computation happens visually with shapes helps in the identification of discrepancies between parts.

Shape grammars can describe with a small set of rules a big set of designs with the same underlying principles and classify the mass of archaeological evidence. Most important, shape grammars as a generative system provides a systematic approach to variation, exploration of all the possibilities within a style, and experimentation and visual testing of different theories, more than a parametric model or schema can do. Knowing the range of possibilities, the historian can make hypothesis, explain and interpret the existence or non-existence of designs among cultures and cultural groups.

The transformations of shape grammars with rule addition, rule deletion or rule change, can produce transformed languages of designs, which is a guide to explore stylistic change and how different styles are related to each other, how one is a transformation of the previous or the later style. Last, the grammar can also incorporate metadata that connect the rules to the documentation of the evidence.

The case study of the ancient libraries was also briefly presented to illustrate the ideas discussed in the paper. It has been argued that the importance of this approach in the analysis of ancient libraries is trifold; first, it explores systematically the range of variant possible scenarios of reconstruction of each library, based on the surviving fragments of the building, as well as the analysis of the other libraries in the corpus; second, it evaluates the validity of the interpretation of debatable buildings as libraries; and third, it predicts the design of other hypothetical but possible libraries that might be excavated one day. A comprehensive account of the grammar of the ancient libraries is currently under way.

References


Grammar Modelling and the Visualisation of an Uncertain Past: the Case of Building 5 at Portus

Matthew Harrison, Simon Keay and Graeme Earl
University of Southampton, UK

Abstract:
The grammar-based procedural modelling technology ‘CGA Shape’ allows the creation of parametric architectural models that can be adapted through both user input and probabilistic processes, and thus can be used to generate alternative visualisations of buildings based on archaeological data. We explore the capacity of this technology to model and communicate the interpretative uncertainty inherent in the process of archaeological visualisation. A review of previous applications of CGA Shape within cultural heritage demonstrates its potential for uncertainty visualisation, but also shows limited incorporation of large, detailed and complex archaeological datasets. The technology is thus assessed through a case study, the reconstruction of a possible ship-shed from the site of Portus, which comprises a heterogeneous dataset based on geophysical, archaeological, architectural and artistic sources. The strengths and limitations of this approach are assessed, considering its influence on interpretative processes, communicative ability, and its potential to provide transparency in the modelling methodology.

Keywords:
Visualisation, Procedural Modelling, Shape Grammars, Uncertainty

1. Introduction

Archaeological datasets are incomplete, ambiguous and continually open to both expansion and revision. Yet visualisations of archaeological sites, structures and objects often fail to communicate this uncertainty to their audience, instead presenting a single holistic image of the past - one that does not differentiate speculation from well-deduced interpretation or hard data. As archaeologists these images are so appealing as research tools due to their ability to crystallise and communicate our interpretation of the past to both academic and lay audiences. But in reality is this an illusion of simplicity that misleads audiences, as well as a practice that limits us as archaeologists to a single path of interpretation?

This paper will explore the use of ‘shape grammar’ based technologies as a process of modelling interpretative uncertainty in archaeology, and to communicate that uncertainty to the public, through the formation of parametric and stochastic procedures to generate multiple visualisations of architecture. Through application of this methodology to an archaeological case study we will test whether the approach can accommodate the interpretative complexities of a dataset based on a wide range of sources, assessing the interpretive influence and transparency of the methodology, as well as the aesthetic and communicative power of the resulting visualisations.

2. Archaeological Visualisation of Uncertainty: Approaches

The issue of certainty and authenticity within digital visualisations of the past is by no means a new one, having been continually raised over the past two decades (Miller and Richards 1995; Ryan 1996; Zuk et al. 2005, Bentkowska-Kafel et al. 2012). Awareness of these issues has caused scholars to suggest the need for new modes of visualisation that signal uncertainty (e.g. Zuk et al. 2005), as well as the documentation of interpretative processes or ‘paradata’ (e.g. Bentkowska-Kafel et al. 2012).

Corresponding author: M.J.Harrison@soton.ac.uk
Despite new methods of visualisation being successfully demonstrated, the vast majority of approaches within cultural heritage have continued unabated to aspire towards the production of photorealistic images that forgo communicating subjectivity to their audience. More than taking focus away from issues of false certainty, the realism afforded by such approaches can be seen to compound the issue; conveying an environment so familiarly ‘real’ to the viewer that one may presuppose an authentic representation of a past reality.

The pursuit of photorealism is in part a perceived need for cultural heritage to keep up with the increasingly impressive images of film, television and gaming. These are partially concerns of public appeal, which are inevitable given the quasi-commercial application of archaeological visualisations in museums and media, but are also pertinent to the communicative power of the image. Through masking the complexity of source data and interpretation, and solely providing the audience with familiar elements of perceived reality (convincing lighting, materials and geometry) as well as established media for representing reality (artistic or photographic composition, lens flares, depth of field, etc.) the image is instantly readable and coherent. It is the consistency and familiarity of the aesthetic composition that make it so legible, i.e. it can be read as a photograph of a past reality, and therein lies the double-edged sword of communicative power and false certainty.

We propose that the academic concerns over the issue of uncertainty and visualisation have made limited impact, despite the validity of their criticisms, because the majority of alternative visualisation approaches compromise one of these two strengths of photorealism: its impressive aesthetic or its communicative simplicity.

Perhaps the most common approach to rectifying issues of uncertainty is to add visual cues within a conventional rendered image to signal levels of confidence or probability of particular elements, such as colour or transparency (for possible indicators see Pang et al. 1997). Yet the mix of familiar and alien aesthetics compromises the coherence of the image. One fails to get an impression of a plausible architectural or artefactual whole when parts of a building or object are perhaps almost invisible or differently coloured, despite getting a clear indication of which parts are most reliably reconstructed. Such approaches have been applied most successfully where there is a single independent factor of uncertainty, such as temporal uncertainty (e.g. Zuk et al. 2005), but how can one accommodate complex dependencies between uncertain elements (i.e. if x is true, y is more likely to be true) as well as communicate a multiplicity of possible forms.

Another approach is to use non-photorealistic rendering techniques to effect a comprehensive aesthetic change, emulating the media of traditional artistic illustration (e.g. Roussou and Drettakis 2003). Replacing photorealistic immersion with a sense of the constructed and subjective through association with traditional illustration acts as a general qualifier to the audience about uncertainty. It is unlikely, however, that such an approach would ever be adopted widely in cultural heritage as the photorealistic is too entrenched in the heritage industry and public expectations.

A seemingly simple solution is to present multiple possible forms of a past environment or object in a series of images. The strength of digital visualisation over traditional illustration has always been the ability to manipulate and alter a model, yet these alterations are still effort intensive. One must therefore consider how many alternatives are reasonable and achievable in terms of labour. In terms of presentation, one must then choose which possible elements to include in final images, in which specific combinations (for example 10 variables each with 10 different possible values would result in ten billion possible final images, even ignoring the possibility of continuous
variables). Furthermore, while each image is in itself coherent, as a whole a collection of images can be overwhelming. As such, frequently one will be highlighted and then reproduced elsewhere as the most probable, as exemplified by the reproduction of one of Millett and James’s alternative illustrations of timber framed buildings at Cowdery’s Down (Millett and James 1983; James 1997, 29).

These limitations of presenting multiple images are remedied by interactive models which utilise parametric relationships. In such systems alternatives can be generated automatically from user input. The implementation of such systems can be seen in early examples with VRML and HTML/java (Roberts and Ryan 1997), as well as bespoke virtual environments such as CREATE (Roussou and Drettakis 2003). In these systems users can explore uncertainty within the parameter-space of a model as defined by the creator, in order to gain an impression of the range of possible pasts without being overwhelmed by several parallel images.

The strengths and limitations of these past approaches to presenting uncertainty in archaeological reconstructions will now be considered in light of the recent development of procedural modelling technologies based on shape grammars, which have been designed for architectural visualisation – namely CGA Shape grammar (Müller et al. 2006a), incorporated into the CityEngine modelling package (Procedural, Esri). This technology affords a system of modelling that can incorporate parameterisation, as well as stochastic or probabilistic expression of these parameters, thereby allowing both dynamic adaptation of geometry as well as quantified conceptualisation of the relative certainty of variables.

3. Shape Grammars and Modelling the Past

3.1 Shape Grammars, Architecture and CGA Shape

Developed in the 1970s as a procedure to generate and conceptualise painting and sculpture (Stiny and Gips 1972), shape grammars consist of a vocabulary of shapes, and rules that successively replace one shape with another to create increasingly complex geometry.

Shape grammars have been used since the 1970s in the analysis and generation of design, including several applications in archaeology and anthropology (e.g. Hodder 1982; Weissman-Knight 1986). More recently, computational procedures based around shape grammars, though incorporating only some aspects of the original concept and adding others, have been developed into commercially available systems to create three-dimensional architectural models - namely CGA Shape grammar (Müller et al. 2006a). Within CGA Shape the rules of replacement predominantly takes the form of subdivision of shapes, and as such is considered a ‘split grammar’ at its core. The underlying premise is that architecture can be expressed as a nested series of such divisions.
into a number of functional or decorative elements (Fig. 1).

Such division-based rules are seemingly evident in the architecture Classical world, but only become explicit in architectural thought during the Renaissance within the re-interpretation and adaptation of Classical architectural style, most notably the work of Palladio (1570).

Within CGA Shape grammar, procedures or rules are written (in an ASCII scripting language) that express the relationship between architectural elements, comprising an adaptable method of geometry creation that and defines the limits within which any example of an architectural schema must fall. The adaptability, and therefore the expression of uncertainty, resides in two elements of the procedure: parameterisation and stochasticism.

Where rules are used to express an architectural form or style, parameters referenced in these rules control the dimensions, angle, colour, material, presence/absence, etc. of specific elements. Control of these parameters allows adaptive, holistic changes that can be used to create variations on a style instantly.

The expression of parameters as either discrete probability distributions (Fig. 2) or uniform continuous probability distributions (Fig. 3) means that the execution of the grammar results in unpredictable variations in the resulting geometry, illustrated in Figure 4. Utilisation of stochasticism to automatically generate variation is at the heart of many procedural modelling technologies from Perlin’s Noise (Perlin 1985) to L-system vegetation (Rusunkiewicz et al. 1996). In these examples the variation is added to phenomena repeated across a modelled scene – a surface/material in the case of Perlin’s Noise and instances of modelled plants in the case of L-systems – in order to mimic the heterogeneity of the physical world without the need for excessive manual input.

3.2 CGA Shape Cultural Heritage Applications

The capability for stochasticism within CGA Shape has been utilised within cultural heritage in much the same way as the aforementioned technologies. Reconstruction of past cities can be achieved by creating rule sets applicable to domestic architectural styles, then generating numerous different examples of that style. Examples include models of Pompeii (Müllerat al 2005), Imperial Rome (Dylla et al. 2010; Wells et al. 2010) and Mayan Xkíché (Müller et al. 2006).
The models produced as part of these projects all have the ability to reflect interpretative uncertainty and to explore hypotheses through the production multiple alternative variations of a building, either by stochastic parameters or simply by user control. Such potential is illustrated by using two of the project rule sets (Pompeii and Rome) to probabilistically generate alternative reconstructions in a paper by Haeglar, Müller and Van Gool (2009). Yet elsewhere the presentation of these projects focusses on single, static, large-scale scenes, particularly when presenting to the public (e.g. Rome Reborn, http://www.romereborn.virginia.edu). More importantly, the aforementioned demonstrations of uncertainty visualisation remain largely hypothetical as they do not accurately reflect specific archaeological data.

The limited incorporation of archaeological data from the site being represented is also evident in the city-scale models of Pompeii (Müller et al. 2005, also distributed with CityEngine) and Rome (Dylla et al. 2010). In the case of Pompeii, the location of each substructure (i.e. room) is derived from the excavated town plan, as is its general type - broadly divided into either atria/peristyles or generic domestic units which may stochastically include elements of shop fronts, grand house entrances, first storey apartments, etc. Beyond this, generalisations of the archaeological data are used to influence the parameters in the form of maps of population density as well as the inferred date and affluence of structures.

All the modelled elements are archaeologically attested from Pompeii or comparable sites, but specific structural data (e.g. the colour of plaster recovered from a substructure, stairs indicating an upper apartment, finds/elements indicating a specific commercial function, known locations of doors and windows, etc.) as well as other key aspects of the interpretation (e.g. relationships between substructures to form coherent domiciles) are seemingly not considered. The model simply reconstructs ‘Pompeiian-style housing’ on a large scale, with extremely broad structure types placed in appropriate locations. In other words, some elements or variables which are known with certainty due to archaeological evidence will not necessarily appear in the generated scene, nor may they be any more likely to appear than alternatives. This is perhaps unsurprising as taking into account all the relevant data for a site such as Pompeii, with both an exceptionally vast and disparate dataset, would be a colossal task.

Similarly, the Rome Reborn project (Dylla et al. 2010) primarily uses CGA Shape to create facades of a known architectural style - the Roman ‘insula’ or apartment block. The volumes
that these facades are applied to are derived from a laser scan of the Gismondi plastico of Imperial Rome created in the mid 20th century. The stochastic rule set creates varied facades that conform to our understanding of that generic architectural form (based on examples from Rome and nearby Ostia) but consideration of historical or archaeological data relevant to a specific structure is absent, or at least is a static reflection of Gismondi’s interpretation of the location and general form of domestic architecture.

In sum, the grammar in these examples becomes a tool to create the effect of an authentic historical city. Although informed in a general sense by archaeological and historical data of architectural styles it is not arguably being used to reconstruct the respective sites: it does not attempt to reflect the available archaeological data for specific structures, its most current interpretation by experts, and associated certainty experts can ascribe to that interpretation.

This is not true of the CGA Shape modelling of Puuc architecture at the site of Xkipché (Müller et al. 2006b), which follows archaeological remains very closely. However it does not make use of random variation in the grammar, undoubtedly due to the high degree of simplicity and formalism of this style of architecture, as well as the exceptional preservation of the buildings on site. The range of forms existent on the site are presented as more or less certain, with the utility of the approach demonstrated to be the economy of creating different forms through alteration of control parameters. The missed opportunity to confront the issue of uncertainty is demonstrated by the authors’ revelation that the buildings may have been plastered and painted with bright colours, though “since their exact appearance is still under archaeological debate, we did not include colorful textures such as paintings in the current model” (Müller et al. 2006b, 145).

A different approach is also evident in the use of CGA Shape for digital reconstruction of the Louvre (Calogero and Arnold 2011). Here two models are created - one is the design of the East Wing that survives today, and the other is a 1663 proposed design that never was realised. The source material for the latter was limited, as only a single ground floor plan and a section through the wing survives. These were supplemented by sketches and an elevation that reconstructed the design of the façade, produced in the 1960s. While there is scope to use the technology to explore the variation possible interpretations of the fragmentary design information, only a single visualisation is produced, reproducing (we assume) the interpretation of the 1960s reconstruction. So while the emphasis is not on economy of labour, it is by no means fully exploring uncertainty. Rather, in this case CGA Shape is being used to aid in the interpretation of the structure and meaning of the two designs, in much the same manner conceived by Stiny and Gipps, an approach also seen in the modelling of Regency architecture in Brighton (Calogero et al. 2011).

With this borne in mind, the illustrations of CGA Shape’s applicability to uncertainty visualisation (in Haeglar, Müller and Van Gool 2009) are problematic in that it has not been established that the grammar can reflect a real, complex, heterogenous dataset with high levels of uncertainty. As experiments they can afford to choose the simplicity of their hypothetical source data, and it is unsurprising that they choose a high level of certainty for high-level variables such as the building scope and basic structure, and add uncertainty on low-level variables such as materials and decorative elements – such as is easily accommodated within the sequential nature of geometry creation in the grammars. In reality, the archaeological record may present data pertinent to a mixture of high or low level variables, often sporadically and in complex, interconnected ways. It remains to be seen in either of these case studies whether the grammar could account for this – accommodating fixed locations and forms of known smaller elements...
within dynamic higher-level changes. It should be noted, however, that several developments of CGA Shape may go some way to solve such issues, though they have not necessarily been incorporated into that available commercially – e.g. ‘snapping’ lines (Müller et al. 2006a, 618) and the persistent naming and modification of elements (Lipp et al. 2008).

The following section will therefore illustrate the authors’ own attempt to use CGA Shape to create a model that reflects interpretative uncertainty based on a real and heterogeneous archaeological dataset.

4. Modelling Building 5 at Portus

4.1 The Site of Portus

Portus is located on the ancient coastline of Italy near the mouth of the Tiber, a few kilometres to the north of the river port of Ostia. The site functioned as a principal port of Rome from the first century AD, its foundation attributed to the Emperor Claudius in around 42AD (Keay et al. 2005).

From the late 1990s the Soprintendenza per i Beni Archaeologici di Ostia has been undertaking a major scheme of conservation and investigation of the site, including excavation. As part of this scheme the Soprintendenza collaborated with the Universities of Southampton, Cambridge and Durham and the British School at Rome in a project of extensive non-destructive survey techniques – both topographic and geophysical. This project ran from 1997-2005, and was undertaken as part of a wider project studying Roman urbanism in the Middle and Lower Tiber Valley directed by Martin Millett and Simon Keay, which in turn forms one part of the British School at Rome’s Tiber Valley Project co-ordinated by Helen Patterson (Keay et al. 2005).

2007 marked the start of a new collaborative project undertaken by the University of Southampton, in conjunction with the British School at Rome, the University of Cambridge and the Soprintendenza Speciale per i Beni Archeologici di Roma, directed by Professor Simon Keay and funded by the Arts and Humanities Research Council. The project, known simply as ‘The Portus Project’, is on-going at the time of publication (see www.portusproject.org, www.heritageportal.eu, and Keay et al. 2009). The project has built on the previous topographic and geophysical survey as well as starting a series of excavations.

4.2 Building 5 – Data Sources and Modelling Methodology

The excavation of a large structure referred to as ‘building 5’, which is located between the site’s two ancient harbours and believed to be a ship-shed or navalia, provides an interesting case study for grammar based modelling. The interpretation of this building relies on different kinds of source data – geophysical survey, excavation, standing remains, as well as drawing on artistic and architectural parallels. Its materials and principal system of structural support is known, being formed of rows of brick-faced concrete piers, yet there is much uncertainty about the structure due both the lack of survival of the upper structure as well as the limited amount of excavation in comparison to the huge scale of the structure – thought to be 60m x 180m (revisions discussed in 4.3).

The excavations have focussed on a small area in the southwest corner of the building. These revealed two internal spaces measuring c.4m and c.12m in width. Standing remains on the site, preserved in a later city wall, show that such units of standard widths are repeated elsewhere, however there are large parts of the structure with no recorded standing remains. To supplement this data a magnetometry survey across the building’s full extent has revealed indications of unexcavated piers. Both the magnetometry and standing remains survey are...
complicated by the fact that the building had a long history with later phases of adaptation (not considered in this reconstruction), so in each case it cannot be ascertained with certainty that indications of structures do not relate to other phases or structures. Taking these limitations into account, together the sources suggested three probable layouts, each consisting of different sequences of the same units (Fig. 5).

To cover the spaces between the piers, a number of alternatives were possible within the limits of Roman engineering. The space could have been covered by a concrete vault, either cross or barrel vaulted, following ‘rules of thumb’ determined by Lancaster’s survey of concrete vaulting in Imperial monuments (Lancaster 2005). It is also possible that the piers supported a timber framed roof, as records of timber lengths and other known timber-roofed structures suggest that beams could be up to 25m in length (Ulrich 2007).

To address the issue of the upper structure, as well as other decorative and structural elements, evidence from other sites was considered. Unfortunately, there are no definitively identified examples of standing or excavated Roman ship-sheds. Therefore the reconstruction drew upon earlier examples from Greek and Punic contexts, as well as later, medieval, examples such as those at Venice, Pisa and Seville (Fig. 6). Representations in Roman art of ship-sheds were also considered. Such comparisons lead to the consideration that the building may have been inclined for the purpose of launching ships, was almost certainly sheltered by a series of pitched roofs, and may have contained various decorative elements.
Figure 7. Twelve stochastically generated models showing alternative reconstructions of building 5 at Portus (exterior), modelled in CityEngine and rendered with mentalray.

Figure 8. Twelve stochastically generated models showing alternative reconstructions of the building 5 at Portus (interior), modelled in CityEngine and rendered with mentalray.
Taking these sources into account rules were created CGA Shape/CityEngine that reflected principles of Roman architecture of this scale, material and function. Over thirty parameters were used to give control of various aspects of the form. These parameters were expressed as either constants, ranges (continuous uniform probability distributions) or probabilities (discrete probability distributions), depending on their certainty and relevant source data (the parameters and source data are summarised in tables 1-3).
The quantification of the level of certainty posed a major problem to this methodology. Figures were determined subjectively in discussion with members of the Portus Project in a fairly informal manner. Whilst archaeologists are familiar with expressing their levels of confidence, it seemed obvious that the way in which individuals quantified it could vary significantly. A more rigorous approach would perhaps be the collation and amalgamation of expert confidences with Bayesian statistical approaches (e.g. Van Leusen et al. 2010, 127). However, this poses some logistical problems in presenting the full dataset necessary to make an informed decision to significant numbers of experts. It also compromises the contextual and temporal nature of the interpretations, i.e. the ways in which they developed, informed one another, and were transformed through the visualisation process. Separating the processes of modelling and certainty quantification would provide a false distinction which could risk limiting a core value of the process – stimulating reflection.

The resulting stochastic procedure was executed twelve times to generate unpredictable combinations of variables. These were then exported from CityEngine and rendered with mentalray using Autodesk 3ds Max (Figs 7 and 8). The resulting images as a whole reflect the interpretative confidence in the various elements of the reconstruction, as well as testing how robust the rule set is in creating plausible forms whilst incorporating uncertainty at various levels of the procedure.

4.3 Assessment

How appropriate is CGA Shape as the basis for a methodology of archaeological reconstruction that foregrounds uncertainty?

4.3.1 CGA Shape as a modelling technology and interpretive tool

To understand CGA Shape’s strengths and limitations one must consider the grammar
as a modelling technology, in the wider sense of being a simplified representation of a real-world phenomenon or system for the purposes of comprehension or testing. Working with a system of rules and parameters, rather than controlling geometry more directly, allows the creator to both conceptualise and document an explicit schema of the building’s organisation, one written in a simple ASCII scripting language. What may be unconscious decisions that go into constructing a visualisation in a system such as 3ds Max or Maya, even if multiple visualisations are being produced, are brought to light through the process of rule-building and parameterisation. Grammar-based modelling is thus a more transparent process, as it presents assumptions about a structure’s organisation (within rules), and identifies key variables along with an assessment of their certainty (within parameters) to other archaeologists and 3D artists. Along with commenting within the code and the documentation of sources informing each parameter, this is a process of providing paradata.

The grammar methodology has acted to guide interpretation through exploration of the parameters, testing hypotheses by assessing the plausibility of overall forms when parameters are pushed to the limits dictated by the source data. It has also helped define where gaps in our data are most crucial, thereby setting new research agendas. Fieldwork at Portus subsequent to the creation of the model has involved full and detailed recording of standing remains in the vicinity of the building 5, primarily in order to reconcile problems in the visualised building layouts.

The methodology’s strength as an adaptable mode of testing hypotheses should be reflected not only by the extent to which it can influence data collection goals, but also the extent to which this new data can continue to be fed back into the model as part of an ongoing interpretive process. Thus far we have only assessed whether the grammar can take into account the complexities of the dataset at the point the model was first created (August 2011), this is the data reflected in tables 1-3 and figures 7-8. However, the aforementioned survey of standing remains, as well as ongoing excavation, has considerably altered our understanding of the structure.

Some aspects of this new data can be accommodated easily within the model, as they relate directly existent parameters. For example, standing remains indicate that the building is longer than first anticipated – a total of 247m – data that can be added in seconds to the appropriate parameter to update the model. However, this increased scope has implications for the layout of the building, as does the discovery of wider internal spaces of c.19m (perhaps functionally distinct), which would require alteration of rules to generate appropriate forms. Deeper excavations have provided new data regarding the inclination of the floor, which is importantly linked to the proposed function of housing ships. The model accommodates both sloped and flat structures, and excavation has indeed suggested a floor which is sloped (though shallower than suggested by Greek parallels at 6.99%-8.75% or 4-5°), however it is now apparent that the foundation levels of piers are constant. It had been assumed in the model that the foundation and floor levels would be concomitant, and this kind of revision of the structural principles of the building will again require revision of the rule-set.
In sum, the extent to which the grammar-based model can only accommodate new data is dependent on how far its creator has considered variability in each aspect of the model, and reflected this in the process of parameterisation. This reminds us that while parametric systems may be more economic ways of creating variations of a visualised form, they do not completely remove concerns of labour economy. Accommodating variation takes effort, both in research and in maintaining a coherent resulting form despite the many potentially conflicting parameters in the procedure. There will always be assumptions inherent in the rule set, and there will always be archaeological discoveries that challenge those assumptions.

4.3.2 CGA Shape as a visualisation technology and communicative tool

As a system of visualisation of uncertainty there are different issues. Our appraisal of past approaches to the issue (above) has suggested that they often compromise the aesthetic of photo-realistic approaches. The capabilities within CityEngine for materials and UVW mapping are limited compared to industry standards. More importantly, the system of geometry creation is focussed on the subdivision of rectilinear faces to create semantically distinct units, which has two important implications. Firstly, it means vertices and edges are placed according to semantic or schematic purposes, whereas in a mesh modelling package they are placed to influence to the resulting aesthetic (particularly important for the purposes of smoothing meshes) and ease of manipulation of the form. Secondly, it means that curved geometry cannot be manipulated within the system easily, being formed of many flat faces. Such complex geometry is incorporated into CityEngine models as ‘assets’ defined in other software used to replace the scope of a rectilinear three-dimensional shape in the model. This presents problems for both the scaling of curved elements (e.g. arch voussoirs) and the interaction of curved elements (e.g. the creation of cross vaults from the intersection of two variable barrel-like intradoses).

The aesthetic shortcomings of the visualisations meant that for the purposes of publicity and publication it was decided that a single model, that deemed the most plausible, would be improved within Autodesk 3ds Max. Due to the aforementioned eccentricity of the mesh created by the grammar based modelling, this was used more as a guide than a starting state, much of the form had to be re-created manually (Fig. 9).

In this sense, the methodology has fallen victim to one of the pitfalls of previous approaches to uncertainty in reconstructions – a single interpretation often takes precedence either for reasons of labour, aesthetics or communicative simplicity. The only output from the methodology that could have been presented to the public was a series of static renders, which as discussed can be overwhelming to the audience. For the potential of the system to be fully realised one must translate its interactivity into a form that can be presented to the public, allowing them to explore the different possible interpretations both through stochastic re-generation of models based on expert assessments, as well as direct control of parameters to allow their own interpretations to be explored within limits set by the creator.

5. Conclusions and Future Directions

CGA Shape and the CityEngine modelling package comprise a tool to model our interpretation of past buildings, incorporating an explicit and transparent statement of the underlying assumptions, alternative forms, and the confidence attached to various elements of that model. It forces one to both confront uncertainty as well as to quantify it. It has a profound effect on interpretive processes, functioning as an adaptable model to test on-going hypotheses and formally identifying the most profound lacunae in the dataset.
Yet as a system of visualisation, from an aesthetic point of view, it cannot compete with the photorealism that dominates cultural heritage, at least on its own. Its strength in terms of modelling uncertainty is its dynamic exploration of innumerable possibilities, and to present a viable alternative to conventional approaches this dynamic nature must be present not only within the modelling stage but also in public presentation.

Developing such an interactive presentation of this structure is an aim of the Portus Project going forward, as is the exploration of newly developed grammar-based modelling systems to overcome the methodological limitations encountered thus far. We have also explored coupling this approach with structural analysis to test the validity of generated structures (Miles et al. this volume), and aim to use these results to feed back into the continuing grammar modelling process.

References


Can Infovis Tools Support the Analysis of Spatio-Temporal Diffusion Patterns in Historic Architecture?

Jean-Yves Blaise and Iwona Dudek
CNRS, France

Abstract:
This paper’s main claim is that analytical reasoning on spatio-temporal diffusion patterns requires a step into abstraction that traditional figurative solutions like maps or 3D virtual models do not encourage. Accordingly, we investigate alternative research practices, namely infovis (Information visualisation) and visual analytics, where the focus is put on revealing patterns of change, and more generally on gaining insight on individuals and collections through visual means. We introduce four graphic combinations implemented on a test case – Zbigniew Dmochowski’s architecture of Poland, a respected and comprehensive classification of architectural facts & trends in Poland over a millennium that combines morphological, stylistic and functional division lines. The contribution presents the pluses and minuses of these combinations, the arguments behind their making, how they have shed (or not) a new light on the test case.

Keywords:
Architectural Heritage, Information Visualisation, Spatio-Temporal Data, Classification

1. Introduction

There are books analysts can rely on in order to analyse pieces of historic architecture in space (dictionaries, maps, etc.) or to spot the time slots they can be connected with (literature about History and History of Architecture). There are also some books that relate changes in shapes over time to a context, and thereby provide means for relations assessment: a typical example is Viollet Le Duc’s encyclopaedia of architecture.

But in all three cases, graphics used almost never fully integrate the time parameter and the spatial features. They rather put them side by side, like in the classic timeline+ cartography paradigm, or provide low-level indications (a limited number of variables taken into consideration) like in the typical map+arrows or map+layers paradigms. Now maybe this was before the computer age? Well, of course not. What computer solutions can we today rely on if we need to understand and represent the patterns of diffusion of an architectural trend, in time and space? Animation techniques, where time is mapped by time? These techniques are well suited for following movements, but they also have proven inefficient in supporting notably explorative tasks.

In this contribution, we investigate how solutions stemming from the fields of Information visualisation (infovis) can reveal patterns of change, and more generally can support background tasks researchers need to carry out when analysing the evolution of historic architecture (relate findings on individual cases to general knowledge, underline contradictions, local architectural inventions, foster the understanding of the individual’s position with regards to trends at that time and that place, renew the interpretation of a stylistic classification; etc).

In a state of the art preliminary section, the contribution positions the above infovis field in terms of scientific legacies, and with regards to spatio-temporal datasets. In this same section we also shortly comment on the visual solutions that will be tested and then focus on the four combinations (Fig. 1) that we have implemented on a test case – Z.Dmochowski’s architecture of Poland (Dmochowski 1956).
However it has to be said that we shall make no general claim on the history of Polish architecture: Dmochowski’s classification – although interesting by many aspects – relies on a subjective, partial selection of edifices. Our intent is to use his classification as a mean to investigate potential benefits of Infovis solutions:

- in uncovering unsaid biases,
- in uncovering spatial and temporal patterns,
- in underlining exceptions, unexpected behaviours,
- in supporting lacking / uncertain / imprecise data.

In the following sections we first come back to Bertin’s vision of “graphics as an answer to a question” (Bertin 2005) and propose a short introduction to Z. Dmochowski’s choices so as to underline what we wish to understand. We then present the pluses and minuses of our combinations and how they have shed (or not) a new light on the test case, before concluding on how the specificity of the data handled in historic sciences can be taken into consideration in the visual analytics process.

2. Conceptual Background

In his clear-cut “brief history of data visualization” Michael Friendly (Friendly 2006) identifies cartography and statistics as the two main scientific legacies of infovis, quoting for instance Galton’s time series or Minard’s figurative maps. By choosing these examples, Michael Friendly somehow already makes the point we wish to make: at the origin of infovis is a concern for spatio-temporal data sets, a need to foster a better understanding of spatio-temporal patterns, and at the end of the day the idea that graphics can help us conduct reasoning tasks on such data sets so as to uncover unthought-of explanations. All through E.R Tufte’s stunning written works further proofs are given that graphics can usefully back up analyses of trends in space and time (Tufte 2001; Tufte 1990; Tufte 1997; Tufte 2006). A number of ground-breaking contributions still emerge today on applications of infovis and related fields to spatio-temporal datasets, as exemplified in (Keim et al. 2011).

In the typical contexts of geovisualisation or GeoSpatial Visual Analytics fields, applications range for instance from analyses of moving objects (Biadgilgn et al. 2011) (Zhao et al. 2008) (Kapler and Wright 2005) – a revival of the time geography paradigm (Chardonnel 2007) - to more abstract developments like the use of visual metaphors for the assessment of semantic relations (Sabol and Scharl 2008).

In parallel, a number of contributions focus on the temporal aspects – may it be in terms of visual solution like (Havre et al. 2008).
2002)’s Theme River, or in terms of time-oriented data modelling challenges like (Aigner et al. 2008). Issues related to the production and readability of graphics – in the shadow of Bertin’s graphic semiology- are also addressed such as the classic 2D vs. 3D debate (Vrotsou et al. 2010).

Yet it has to be said that infovis covers a wide range of application fields, far beyond spatio-temporal data sets – and a comprehensive overview of related works would be, if not unreachable, for sure irrelevant here. Moreover, the term infovis itself has here and there been complemented or replaced by other terms such as Knowledge Visualisation or Visual analytics. So we consider more fruitful to briefly introduce in the next subsection some definitions and pinpoint some differences, before coming back to this contribution’s storyline.

2.1 Infovis and Visual Analytics

To start with a definition has to be given for the term “visualisation”, a term that today can be found in literature as a sort of flabby synonym for representation, and used in the context of numerous practices (abstract diagrams, cartography, CAD and GIS, realistic rendering, etc.). As far as information and data visualisation are concerned, visualisation is not about drawing but about thinking – in the words of (Card et al 1999) it is about using vision to think. In his overview of the infovis field, R.Spence (Spence 2001) defines visualisation as a cognitive activity, underlining the idea that we rephrased in the informative modelling 15th rule: If your graphics did not lead to new insight on your data, consider it useless (Dudek and Blaise 2006). In other words, whereas representation is an end – the end of a cognitive process; visualisation is a mean – a mean to perform reasoning tasks all along that process.

Infovis is defined by W. Kienreich (Kienreich 2006) as the use of computer-supported, interactive, visual representations of abstract data to amplify cognition. W. Kienreich shows that an infovis solution is most often composed of three fundamental units combined inside a master visualisation: visual formalisms (using a diagrammatic visual language to convey info in an abstract way – bar charts, timelines, histograms - see Fig. 2), visual metaphors (basing on real world equivalents – a city layout for a bibliography for instance), and visual models (in cases when the information is itself a real-world equivalent – typically cartography).

On the y axis, we represent the overall amount of changes and alternatives for the whole collection of 27 artefacts under scrutiny, counted every fifty years.

Reddish rectangles represent, from bottom to top, morphological changes (a storey added to the artefact for instances), episodic changes (basically maintenance, however can include major changes that have no impact on

Figure 2. Represented as a classic histogram, a time series (50 years granularity) showing density of changes on Kraków’s market square (Dudek and Blaise 2011).
the artefact’s morphology such as change of roof material), destructive changes (partial or total destruction).

Greenish rectangles represent confirmed (bottom) and unconfirmed (top, darker) alternatives. Confirmed alternatives mean we deal with possible divergent options on what did occur, basing on indications that are duly established and related to the artefact under scrutiny. Unconfirmed alternatives correspond to indications that may be taken into consideration or may be not – because the information is questionable, or more often because it is not directly related to the artefact under scrutiny but to a neighbour, or any semantic group it may belong to (typically useful when considering possible (but unaccounted for) consequences of a fire on a neighbouring artefact).

In this visualisation note for instance for the first half of the XIXth century (1807-1857) that episodic changes (modifications - orange) widely outnumber destructive changes (brownish red). The visualisation shows the “pattern of destruction” usually associated with this period should rather be called “pattern of renunciation” in front of maintenance and repair difficulties.

Infovis is an expanding field, with a number of branches we are not particularly interested in (in the context of this contribution). What makes an infovis solution distinguishable is the info (abstract information, large and/or complex information spaces) and the services expected (information seeking tasks, information discovery tasks). By contrast, the sometimes disputed field of knowledge visualisation often works on smaller, but highly organized sets of information, and focuses on the transfer of knowledge among persons – may the information be abstract or not (Kienreich 2006).

Finally, the recent field of visual analytics is defined as an outgrowth of the fields of information visualization and scientific visualization: it focuses on “analytical reasoning facilitated by interactive visual interfaces” (Thomas and Cook 2005). Obviously the dispute on “which is which” is of poor relevance here. What is key to mention though is that these fields share on one hand overlapping legacies and on the other hand common techniques, goals, and sometimes methods.

2.2 The Visual Solutions Tested

Over the years, a large number of visual solutions have been introduced to cope with spatio-temporal or time oriented data sets – an overview of this latter topic can be found in (Aigner et al. 2011). In this paper we shall analyse combinations of the following:

Timeline (Fig. 3a) this is a concept we do not need to comment on very long, as its use is rather common (in particular in Western cultures where time is seen as linear). Naturally this is its major strong point: it is easily understood. In fact when analysing it from closer a timeline is strongly restrictive: it implies a de-facto discrete time model, needs continuity in the time scale to be readable, it is poor in assessing rhythms and frequencies, and furthermore it better adapted to long walls than to screens where it requires tedious zoom and pan interactions (see on this Mostern and Johnson 2008 or Blaise and Dudek 2008).

Small multiples (Fig. 3b): this expression was introduced by E.R. Tuft (Tuft, 1990), but the concept itself is attributed by M. Friendly (Friendly, 2006) to Christopher Scheiner’s 1611 solution to show the changing configuration of sunspots over time. The idea is to show in a grid-like layout successive spatial distributions of this or that object, of this or that phenomenon, thereby allowing a visual analysis of changes. This solution was for instance used by Galton in his XIXth century multivariate weather charts to uncover meteorological trends (Friendly, 2006), and applied on historical data sets in (Dudek and Blaise 2011).
Concentric time (Fig. 3c): this is a visualisation we introduced in a recent experiment on the development of the market square in Kraków, that combines time and space. It is composed of a central 2D map, with all around it expanding concentric circles that represent a move towards the past. To each artefact positioned inside the 2D map corresponds a classic timeline, developing radially, along which events connected to the artefact’s evolution can be localised in time. Various interactions between the temporal indications along the timelines and the spatial indications inside the central 2D map are then implemented to facilitate the analysis (Blaise and Dudek 2011).

Multidimensional icons (Fig. 3d): there are a number of striking examples of how multidimensional icons can help fostering insight “at a glance” in (Spence 2001) or in (McCandless 2010). The idea is to summarise the values of several variables for a given item in a collection through a simple 2D icon – often relying on analogies or metaphors. In the context of spatio-temporal datasets this solution however has drawbacks. Multidimensional icons are abstract representations (spatial features are hard to convey), and the number of variables that can be compared, in a readable way, is an issue (Chan 2006). Multidimensional icons are usually extremely efficient in depicting an individual item’s feature – they are less efficient in underling similarities and trends.

TimeWheel (Fig. 3e): this concept, first introduced in (Tominski et al, 2004), combines a classic horizontal timeline with segments distributed around the timeline in the shape of a polygon. Each segment of the polygon represents the range of values of a given attribute: for instance a length, an age, a price, etc. An item is localised in time on the central timeline, and then lines connect it to values of its attributes on each segment of the polygon. Although at first glance it may appear complex, the concept is definitely efficient in underlining trends inside a collection. In return, one to one comparisons are not that readable and naturally spatial distribution is clearly hard to convey.

3. The Test Case

In the introduction of his historical survey, Z.Dmochowski, a Polish architect exiled in Great Britain during WWII, defines his objective as such: “show the development of architecture within the polish State from the Xth to the XIXth century” (Dmochowski, 1956, pp. xvii). Yet Dmochowski’s book is more than a survey, it is a de-facto classification - yet with little indications on the division lines used. Some hints are given in the introduction: “the book divides the material according to style, but the distinctions should not be
pressed too far”. Dmochowski pushes this doubt further by adding “It is a great help, in studying the history of architecture of any country, to know something of its political [...] background”: a way to put stylistic affiliations into a wider perspective. Finally, his definition of architecture - “[...] architecture, which of all the arts is the most socially conditioned” – further questions classic stylistic division lines. In short, Dmochowski acknowledges that styles are a key marker of trends, but in the same time that there are other aspects – context in a broad sense – that are equally key in a sensemaking classification effort.

Dmochowski’s survey is basically a description and a classification of individual edifices, with varying amounts of information given depending on the edifice. For a majority, elements of chronology and a spatial layout (2D plan) are given. Other pieces of information on the morphology are more unevenly available. Edifices are segregated by groups (one group one chapter) that correspond to the third step in a three-levels classification (Fig. 4):

- At step 1, styles act as the main division line (Romanesque, Gothic, Renaissance, etc.)
- At step 2, use and/or destination of the edifice refine the classification (Romanesque secular architecture, Romanesque ecclesiastical architecture, etc.)
- At step 3 – groups - architectural composition and choices are introduced in order to further segregate the collection (Basilican churches with transept, Basilican churches without transept, Single cell circular churches, etc.).

Beyond these three division steps there are other unsaid steps: space and time. Indications given here by Z.Dmochowski are quite hereterogeneous (dates like “XIIth century, or “1086”) and furthermore this variability differs from group to group.

Now at this stage it has to be made clear that our goal was not to come out with an exhaustive analysis of Dmochowski’s survey, but to conduct something like a like a proof of feasibility experiment. So the experiment we report was carried out on a subset of

---

**Figure 4. Classification levels, groups considered in this experiment, examples of edifices.**
Dmochowski’s survey: 50 ecclesiastic edifices, 2 styles (Romanesque /Gothic) 10 groups (Fig. 4), time and space, plus a subjective selection of features (our interpretation step: dimensional and architectural features, corresponding to heterogeneous variables, not systematically present in the survey). In short, from Dmochowski’s survey we extracted an information set that can appear as rather simple, with eight variables to combine (Fig. 5):

- Style,
- Group,
- Foundation date,
- Localisation (2D coordinates),
- Length,
- Shape of the Apse,
- Number of chapels,
- Number of Interior volumes.

Yet it might not be that simple –even with such a limited test case there are some clear difficulties:

- Unsaid division lines in Dmochowski’s classification,
- Heterogeneous variables,
- Lacking / imprecise pieces of data,
- Handling of BOTH time and space,
- Need to depict each item, its relation to others, and the global features of the collection.

Well these difficulties are precisely what makes his input a great test case: our objective will be to try and see, given these difficulties, whether or not visual tools can help us uncover information and knowledge, whether they can help us gain some insight on his classification effort, on each of the 50 edifices, and on the 50 of them as a collection.

### 3.1 Timeline + Small Multiples

In this first combination a point corresponding to the foundation date of the edifice is localised along a classic timeline. In parallel, inspired by the small multiples precedent, we sum up inside a square the seven remaining variables through colour, value, position, shape. The square’s background colour and its value are used to convey the style and group variables (Fig. 6a). The square acts as an equivalent to the map of Poland – more precisely a map of Poland is inscribed inside the square. As a consequence the localisation (2D geographic coordinates) of the edifice can
be interpreted as \((x,y)\) coordinates inside the square (Fig. 6b). Values for length, number of chapels and number of interior volumes are reported on three axes, forming a triangle (Fig. 6c). Finally, a small icon on the top right corner identifies the shape of the apse (Fig. 6d).

But then how can these squares be connected to a chronology, to the timeline? If we just align them with the timeline there will be either clumsy overlapping when edifices are close in time – or a lengthy visualisation unsuited to a context view (the whole collection within the eyespan). In fact this combination is not well suited to reading the chronology in discrete or continuous time. Instead, we use it to read the chronology in ordinal time (only the order is known): each group is represented with its oldest item duly aligned on the timeline (Fig. 7a), and the rest of the group simply positioned horizontally by order of appearance. Interactions on each small multiple do allow the reader to precisely date it on the timeline (Fig. 7c).

The result is a clear insight gained on the variability of items inside each group, and accordingly on groups consistencies (Fig. 7b). Exceptional behaviours for one specific variable are also underlined (Fig. 7d), as well as cases when the dating is questioned or imprecise (Fig. 7e). A point can for instance be made on this subset of Dmochowski’s survey: the most “recent” group the least accurate in terms of dating...

A projection of each square along the timeline at a position corresponding to the actual foundation dates for each edifice can be added so as to highlight densities of change, or to correlate changes to other factors.

This combination does have its weaknesses: ordinal time is convenient when wanting to reveal group behaviours, but costly in terms of readability for individual items. Furthermore, as it is, from this visualisation only a rather vague understanding of the consistency of a group in space is within reach.

3.2 Concentric Time + Small Multiples

This combination reuses the small multiple introduced above, but with the ambition to better handle and visualise the space parameter. A map of Poland is drawn in the central spatial layout of the visualisation, with coloured dots that position edifices (colour and value are used to convey the style and group variables). In cases when several edifices are localised in the same city rectangles are used instead of a dot. To each edifice also corresponds a dot on the concentric circles – the older the farther from the central spatial layout. Each item in the collection, each edifice, is thereby localised in time and positioned in space. On user demand items can be browsed one by one, opening the small multiple and a plan view. This modality corresponds to a focus view: it allows one to one feature comparisons inside a group, between neighbouring edifices, etc. Alternatively the whole collection of small multiples can be shown together. This modality corresponds to a context view: it delivers an overview of spatial and temporal features across the collection.

At this stage the visualisation does allow a better analysis of spatial features, but groups are not as readable. In addition, by scattering squares all around the visual display, it in fact
has little to do anymore with the very concept of small multiples. This combination can however be noticeably ameliorated when representing a group by an arc located at a position corresponding to the average foundation date for edifices of the group.

When doing so, temporal consistency inside the group is clearly highlighted (Fig. 8a) – a point that could not be made with the previous solution. Furthermore, if interaction is implemented between the arc representing the group and the dots on the map, clear (yet unsaid in Dmochowski’s survey) spatial patterns are revealed. For instance, group 9 “Double nave hall churches” appears typical of Małopolska (south-east, around the former Polish capital Kraków) whereas group 8 “Three nave hall churches” is widespread rather in the northern half of Poland (Fig. 8b).

In short, this solution probably has a higher potential in rendering both spatial and temporal aspects. Yet it implies a steeper learning curve, and needs stronger interactions.

3.3 Cartography + Multidimensional Icons

When handling geo-related pieces of data one of the first reaction analysts tend to have is to produce some kind of cartography. And there is no reason to believe cartography can’t be used in the context of infovis solutions. In this third combination we start from a basic 2D cartography of Poland on which we position edifices as we did in the previous solution (coloured dots). Each edifice is identified by the system as belonging either to the north or to the south part of Poland, and then either to the east or to west part of Poland. As a result each edifice on the map can be connected to two multidimensional icons located on the sides of the map (Fig. 9a). The remaining variables are summarised in a multidimensional “flower-like” icon composed as follows:

- Length of the stalk corresponds to the distance, in time, of the current edifice to the oldest inside the collection (here 990). The longer the stalk, the younger the edifice (Fig. 9b).
- Diameter of the corolla’s centre: length of the edifice (Fig. 9c).
- Number of petals: number of interior volumes (Fig. 9d).
- Shape of bottom leaf: shape of the apse (Fig. 9e).
- Number of leaves: number of chapels (displayed on user demand).

Figure 8. Assessing group behaviours with the concentric time visualisation.

Figure 9. Combining a basic cartography with multidimensional icons.
Ill-defined positions and dating are represented by line styles and colour.

Used in context view (Fig. 10a), the visualisation allows east/west; north/south comparisons. It helps for instance uncovering different distribution patterns between the west (regular distribution in space) and the east (imbalanced distribution in space), or between the south (high diversity in groups present, consistent and long lasting time distribution) and the north (less diversity in groups, smaller temporal coverage).

Used in focus view (Fig. 10b), the visualisation allows item to item comparisons among neighbours and helps for instance reading spatially consistent group behaviours.

On the overall, this solution over-performs the previous ones on one aspect mainly: its capacity to superimpose some layer of spatial analysis – exemplified here by the four sectors tested.

However the solution also has its drawbacks – the spatial analysis layer requires a high level of interaction, and temporal aspects are seemingly rendered with less efficiency. In detail, some choices on the making of the multidimensional icon are questionable – typically representing a 1D value by a 2D object (mapping the length to the diameter of a circle) is a notably risky strategy.

Before developing the fourth solution it is important to make two intermediary statements.

Whereas with cartography one tends to pack everything up inside a unique system of representation – in this approach – infovis applied to historic data sets - we rather make use various systems of representation.

Up to here the three solutions proposed remain item-oriented: each edifice is represented separately, here by a small multiple, there by a multidimensional icon, etc. Accordingly these solutions allow one-to-one comparisons, but are not necessarily well suited to analyses of the collection as a whole. Typically, if my question is “are apses in majority polygonal” the above graphics can give the answer, but with a boresome decoding effort from the user.

3.4 TimeWheel

The TimeWheel concept (Aigner et al, 2008) combines a horizontal timeline with segments distributed it, along which values of various attributes can be reported. In our experiment foundation dates are distributed along the central timeline, the length of which being this of the time interval between the “oldest” and the “youngest” edifice in our collection. The polygon used is a hexagon, with a bottom side used to convey two variables: groups and style, reusing colour and value as for the previous combination (Fig. 12a). On either side of the hexagon are four segments used to convey the variables length, shape of the apse, number of chapels, and number of interior volumes (Fig. 12b). Finally, we chose here to substitute to the geographical coordinates of the edifice another indication:
the size (number of inhabitants) of its host city today (Fig. 12c). At first glance the visualisation may appear rather visually crowded – but with some interaction it reveals its full potential (the time overlapping of styles is for instance clear-cut to read Fig. 11a).

We represent on each segment all the values for the corresponding variable inside the collection by a point (for instance, all lengths are aligned in increasing order). As a result densities are straightforward to read (Fig. 11b).

Finally, this visualisation proves efficient in analysing how the value of a given variable varies over the whole collection. In the example above (number of chapels) value “zero” – no chapels - occurs almost in every group and style (Fig. 12g). By contrast, high values always occur inside the same 2 groups (Fig. 12h).

However, the “spatial distribution” feature is noticeably absent from this visualisation. This calls for a general remark: when dealing with spatio–temporal data sets in the context of history, one probably needs to come out with a specific, fine-tuned blend between abstraction – the infovis legacy - and figuration - the storytelling legacy.

3.5 Interaction and Implementation

In the context of this written contribution we basically show screenshots of a set of computer-supported, interactive graphics. These graphics are created dynamically, at query time, from the reading of an RDBMS system storing the actual data. They are written in the standard SVG (Scalable Vector Graphics) format (ASCII format). The whole process is supported by ad-hoc programming, avoiding dependence on any piece of commercial software. Accordingly the can be easily duplicated, provided some knowledge of RDBMS and basic programming skills are at hand.

Graphics can be read on standard web browsers, and freely distributed. Interactions can be attached to each geometric object inside the SVG file, may it be a simple background line or a group – a multidimensional icon for instance. We do acknowledge that a number of issues are left uncommented in this paper (possible data pre-processing here and there, escaping from two-dimensional visualisations, cognitive load and evaluation in general, etc.). The fact is that they would drive us too far from the paper’s main claim, and require quite extensive (and unplanned by the editors) explanations since visualisations we tried out would each call for specific comments.
4. Conclusions

The first conclusion we draw can be seen as a tribute to Bertin’s view: since efficient graphics are those designed as answers to a question, it is a necessity when facing several questions, to promote several graphics. Indeed, the four solutions we have tested shed different lights on our data set. Accordingly there is nothing like a universal visualisation – there are only visualisations that do help uncovering information, and the others - useless visualisations.

Coming back to the test case, the visual solutions we have proposed did help us point out some interesting figures from Dmochowski’s survey:

- basic spatio-temporal patterns, like a move northbound as time and styles pass by,
- architectural patterns, like higher or lower consistency in the architectural composition features depending on the groups,
- unsaid biases, like a relative (and questionable) homogeneity in the number of items presented in each group.

But it would take more than a sub-set of the survey, more detailed architectural analysis of the edifices chosen, more cross-examination with historic factors, to really question Dmochowski’s classification.

At first glance, what we have proposed in this contribution is a method to re-read, re-interpret existing spatio-temporal data sets. From closer, we have exemplified this approach through a set of graphics that each with its pluses and minuses help uncovering patterns of the underlying data set, and have shed a new light on Dmochowki’s classification effort.

Infovis and its outgrowth visual analytics are precisely about that: re-investigating a data set through visual means, and thereby giving ourselves chances to reach to some new conclusion, or at least to renew the way we picture the data set to ourselves.

However methods, tools and practises from these fields rely on human reasoning capabilities: the graphics by themselves are of no use if:

- They are not the result of adequate modelling and filtering choices (for instance keeping tracks of doubts),
- They are dedicated to communication rather than to perform reasoning tasks.

We hope this contribution showed infovis/visual analytics solutions can apply to test cases like ours, and can be fruitful even when dealing with history-oriented data sets. At the same time it has to be stressed that not all of the infovis/visual analytics methodology applies - solutions from these fields are not solutions unless we fully master the process from data modelling to the production and coding of graphics. In general scientists dealing with history-related data sets should be cautious with statistics-oriented formalisms and/or automatic methods, like for instance clustering (possibly deceptive).

To conclude, we wish to add that infovis solutions should not be seen as yet another technological magic wand – but they can, and probably should, be seen as yet another mean for us to re-examine data sets in the light of our knowledge of historical data sets.

References


History in 3D: New Virtualization Techniques for Innovative Architectural and Archaeological Scholarship and Education

James C. Sweet, Krupali Krusche, Christopher R. Sweet, and Paul Turner

University of Notre Dame, USA

Abstract:
Emerging photographic and 3D scanner technology has provided an unprecedented opportunity to document historical structures in an electronic form that can be easily recorded, searched and disseminated. Manufactures such as Leica (Leica Geosystems 2010) have taken technologies applicable to space exploration and produced scanners with resolutions in the millimeter range and built in photographic capture. GigaPan (GigaPan Systems 2010) photographic technology uses a programmable robotic tripod mount with a Digital Single Lens Reflex (DSLR) camera attached to produce high-resolution panoramic digital images. For the generation of immersive environments, and the ability to make discoveries after the completion of a site visit by virtual exploration, it is desirable to amalgamate these technologies. The authors present a method of texturing scanned 3D “point cloud” data with images obtained from either the scanner’s camera or GigaPan capture in an automated process. This avoids the difficult, and approximate process of manually registering the point cloud with individual images and hence allows much larger models, with many images, to be processed.

Keywords:
3D, Gigapan, Leica, Visarray, Visualization, Virtualization, Textured Models

1. Introduction

New technologies have made it possible to convert the archaeological and architectural records of the past into digital formats and then record that information in databases, Computer Aided Design (CAD) maps and digital images. Data created using state-of-the-art 3D laser scanners and high-resolution panoramic digital images offer scholars the unprecedented opportunity to recreate accurate, searchable digital copies of the sites they are researching. Digital records of historic sites and structures with an accuracy of 4mm potentially gives researchers, educators, students and the general public an opportunity to virtually explore all the details of a site.

Researchers working at archaeological and architectural sites typically produce a diverse array of 3D data allowing them to make more accurate claims about historical sites. This paper presents the results of our research combining 3D scanner and high resolution camera images to create unique data virtualizations, allowing users to frame new research questions and generate innovative ways of interactive explorations of historic properties and world heritage sites. Using data already collected at the Roman Forum, we propose to develop web accessible software capable of fusing extremely high-resolution (giga-pixel) panoramic images produced by the GigaPan system, with dense point clouds generated by Leica 3D scanner to create accurate, interactive 3D images of archaeological sites. Because the technique affords unique ways to manipulate and interpret 3D virtualizations of historic sites, its potential contribution to the field of humanistic scholarship and education is vast.

In addition to offering unique ways to study historic sites, the paper presents methods to spur new scholarship, enabling new methods of analysis, site manipulation and reverse blueprinting (a way of retroactively generating plans should the site ever be destroyed).
Learning materials generated will allow better study of the conservation arts, what it means to think in 3D, and connections across pencil, paint and pixel modelling. Our approach should add an invaluable resource for educators/scholars seeking to enrich their curricula or research agenda with the study of historic sites.

3D visualizations of historic, world heritage and cultural sites will not only expand and enhance our collective understanding of the historical record; it will bring the past to life for a population who is increasingly visually oriented. This precise type of 3D modelling allows humanities students and researchers to pose questions surrounding what it means to practice conservation and preservation of crucial sites in the age of digital virtualization.

In Section 2 we introduce the DHARMA group and the history behind their work at the Roman Forum. In Section 3 we describe the details of scanning the Roman Forum and the workflow employed. In Section 4 we discuss the GigaPan photographic work done at the Roman Forum and in Section 5 we discuss the methods we have developed to combine these data sets. Finally in Section 6 we describe the workflow for data combination followed by our conclusions in Section 7. We also include an Appendix to describe the interchange file format that we have developed and a section on the Leica data scanner.

2. Dharmma - Roman Forum Project

Digital Historical Architectural Research and Material Analysis (DHARMA) is a research team founded in 2007 based at the University of Notre Dame School of Architecture. The team, under the direction of Prof. Krupali Krusche, works on documenting historic monuments and World Heritage Sites around the world with the use of Leica 3D laser scanners. These high-speed, long-range scanners are ideal for projects that are difficult to document by traditional methods. The scanner provides researchers with the most field-efficient means of data collection. Recently the team has also used 3D scanning to assess aging effects on historic buildings and reconstruction processes of buildings with historical value.

In the summer of 2010, the DHARMA team, led by Prof. Krusche in cooperation with Archaeologist James Packer, and with special permissions from Soprintendenza Speciale per i Beni Archeologici di Roma, Ministry of Heritage and Culture and the Archaeological Service, digitally and traditionally documented the Roman Forum site, Rome, Italy in a number of different formats.

3. 3D Scanning

In the recent years, laser scanning or Light Induced Detection and Ranging (LIDAR), has become a major technique that has started replacing the use of hand measuring, surveying, and the use of total station in the field. A laser scanner is an automated surveying apparatus that uses a laser beam to collect location coordinates from the surface of a desired object. These measured coordinates are recorded in the form of a point cloud that identifies the shape of the object by converting the spatial geometry established through the x, y, and z position of that point in space. Depending on the way the data is collected, the scanners have different capacities and are of various types. This paper concentrates on the “time of flight” type of scanner, which is specifically used in the historic and archeological documentation of large sites. The data produced from the scanner can be used to create measured-drawing documentation of historic monuments and existing buildings, to save costs on as-built drawings used in restoration and reconstruction.

The Roman Forum data was collected using hand measuring, photogrammetry, 3D Scanner and GigaPan technologies. A team of two graduate research assistants scanned the Roman Forum site, using the Leica Scanstation model, with a resolution of 1cm x 1cm
throughout the site (Fig. 1a). The team worked on the central Forum i.e., the open space (the “Area Fori” including the east and west rostra, the small monuments around the latter, and the Diocletianic columns) and the following buildings: the Temple of Caesar, the Temple of Antoninus and Faustina, the Basilica Aemilia, the Curia, the Arch of Severus, the Temple of Concord, the Temple of Vespasian, the facade of the Tabularium, the Temple of Saturn, and the Basilica Julia.

All spaces were assessed, analyzed, documented and scanned from outside, as well as inside the ruins. It took the team seven, twelve-hour workdays to complete the project with one scanner and two battery units being charged interchangeably. The site is approximately 150m x 250m in dimension with very large variations in contour, and about fourteen monuments and change of grade existing throughout the area. It took the team 27 scans, with more than 100 million data points of cloud data, and five strategic target locations to completely document the site. A team of four DHARMA undergraduate members documented the whole site with hand measuring techniques. They completed measuring the over-all site with sketch drawings containing dimensions of individual buildings and their location from a set zero point.

Figure 1. (a) Luke Golesh operating the Leica ScanStation at the Roman Forum. This image is taken from top of the ruins of Vespasian temple; (b) The Roman Forum as captured by the Leica ScanStation.

Figure 2. Results of the measurements taken at Temple of Saturn, at the Roman Forum, with the Leica ScanStation (a) and hand measure drawings (b).
While the information collected by hand measuring was adequate to give an overall understanding of individual monument dimensions, a lot of time was spent to maneuvering through the complex site, and the detail of information was not exact or refined because of the ruin state of the site in relation to what was collected by the scanner. It is important to note however, that there were very important observations recorded during the hand surveying of the site that would have been missed while scanning the site from remote locations (Fig. 2a-b). The scanner did exceptionally well in capturing data from points that were not visible or were overseen at the site in many ways due to their remote location and issues of accessibility. Even in places where the scanner couldn’t be positioned because of time, target and location constraints, we were still able to document around 95% of the Forum site by being strategic about scanner positions throughout the field work.

Off-site, the registered and unified data from the 27 scans revealed information that the scanner had collected filling “holes” in the scan results for inaccessible positions. The plan and sectional views of the combined 3D scan data revealed information regarding the spatial correlation of the individual ruins as never seen before. It also revealed the comparative level change of the archaeological site in relation to the present city of Rome (Fig. 1b).

4. Gigapan - Photographic Data Acquisition

Panoramic photo documentation of the Roman Forum was carried out by Ben Keller from the Academic Technologies team under the direction of Paul Turner. Keller accompanied the DHARMA laser scanning team to photograph the site using GigaPan technology originally developed by robotics scientists from Carnegie Mellon University in cooperation with NASA Ames Research Center for use by the Mars Rover program.

The GigaPan system uses a programmable robotic mount to precisely control a Digital Single Lens Reflex (DSLR) camera to take hundreds or thousands of pictures of a scene automatically. The practical resolution limit of the panorama imagery created using GigaPan technology is directly related to the zoom level of the lens and quality of the camera’s image sensor. The resulting digital images are then stitched into extremely high-resolution panoramas via GigaPan’s Stitch or Kolor’s Autopano Pro software. Autopano Pro uses image stitching technology developed as part of a research project at the University of British Columbia in Vancouver (Canada). The resulting panorama images are then uploaded to the web using GigaPan Uploader and displayed via http://www.gigapan.com. GigaPan images can also be embedded into web pages, or displayed via the GigaPan Layer by Google Earth (Google Inc. 2012). Using this technology the visual and scientific off-site study of the Forum monuments was greatly enhanced. GigaPan was extremely useful in documenting small-scale details of many monuments at high heights and provided access to many features not visible to the naked eye from ground level.

Careful attention was devoted in post-site processing of the original individual images captured under the control of the GigaPan robotic mount to combine these large sets of digital images with the least amount of parallax and best exposure levels to create high quality panorama images. After stitching the resulting panorama image files are many giga-bytes and contain billions of pixels that present challenges when using Adobe Photoshop or similar single image editing software products for image enhancement. Figure 3 illustrates an example where 150 separate high resolution images, taken using a DSLR mounted onto the GigaPan robotic tripod, were stitched, using Kolor’s Autopano software, into a single GigaPan panorama image. This can be viewed in a web browser at various zoom levels using image streaming techniques adapted from Google for displaying online maps, photos
and 3D objects. Zooming into small sections of the arch in the GigaPan image reveals fine details not apparent to the naked eye when viewing from ground level.

5. Combining Leica Scan Data and Gigapan Images

The VisArray project team at Notre Dame’s Center for Research Computing has been exploring methods for combining the two primary technologies into a single interactive system for cataloging, analyzing and displaying both point and raster data at both standard and extremely high resolution.

Previous methods have manually adjusted the registration between the data sets visually, which does not guarantee the fidelity of registration over the full extents of the data sets and, in general, cannot correctly align the data sets. In addition, for many images an automatic process is desirable. Computer programs also exist to approximately align the data to produce single point visualizations, in general this visualization cannot be rotated through all viewpoints and it is not possible to measure surface features based on the underlying 3D data.

In this work, 3D data is combined with 2D surface data such that:

1. The registration between the data sets, and hence the validity of observations and measurements based on them, is determined geometrically and maintained over the surface of interest.

2. Where multiple sets of 2D data are available, the optimal set can be determined for each rendered pixel set. In addition, visual artifacts generally present with multiple images can be reduced.

3. The surface images do not have to be produced from fixed viewpoints in relation to the 3D representation (or map directly to individual 3D features of the object), allowing the acquisition of data from the architectural to nano scales. This is particularly important for complex surfaces.

To accomplish this, we have a main source of point data (a Leica scanner), and two potential sources of image data (a Leica scanner or a DSLR attached to a GigaPan). To work with the Leica scanner’s picture data, we had to understand the camera’s motion during scanning. We also had to export data from Cyclone (Leica’s scanner software) that describes some extra information about the camera’s specifications. Finally, we determined that the camera data is arranged in a gimbal format i.e. z,y,x rotation order.

To extend this to using GigaPan picture mapping, it becomes more complicated. This is because, unlike the Leica scanner, the GigaPan will not be able to take images from the exact same location as the scanner. Also, due to the nature of GigaPan, end users attach a DSLR to the device and each of these will affect the field of view and stepping, which are important parameters to mapping the pictures onto the data.
def CalculateSurface(Horizontal, Vertical, Points):
    Vertices = []
    for i in range(Horizontal - 1):
        for j in range(Vertical - 1):
            lineA = i * Vertical + j; lineB = (i+1) * Vertical + j
            a = Points.get(lineA); b = Points.get(lineB)
            c = Points.get(lineB + 1); d = Points.get(lineA + 1)

            count = 4
            if(a.isZero) count -= 1
            if(b.isZero) count -= 1
            if(c.isZero) count -= 1
            if(d.isZero) count -= 1

            if(count < 3) continue

            if(a.isZero):
                Vertices.append(b); Vertices.append(c); Vertices.append(d)
                continue
            if(b.isZero):
                Vertices.append(a); Vertices.append(c); Vertices.append(d)
                continue
            if(c.isZero):
                Vertices.append(a); Vertices.append(b); Vertices.append(d)
                continue
            if(d.isZero):
                Vertices.append(a); Vertices.append(b); Vertices.append(c)
                continue

    Vertices.append(a); Vertices.append(b); Vertices.append(c)
    Vertices.append(c); Vertices.append(d); Vertices.append(a)
    return Vertices

Listing 1. Point to Surface Mapping.
5.1 VisArray Viewer

The software developed for this application is based on the VisArray framework (http://www.visarray.com) for the Java, iPhone/iPad and Android platforms.

VisArray is a visualization framework for rendering scientific data. By combining data acquisition and manipulation in front-end modules with flexible back-end rendering modules our framework has been targeted at a diverse range of projects, from clustering of social network data to visualizing fibrin networks formed during thrombus development. The resulting display can be rendered to traditional displays, 3D hardware, ray traced, embedded in web-portals and, more recently, projected onto modern Digital Visualization Theater facilities.

The framework provides a front-end module which encapsulates the view manipulation, specific rendering methods and Graphical User Interface (GUI) elements of the viewer. The application specific code then inherits this module to produce a viewer tailored to the collaborators needs with minimal development. One or more output modules can then be chosen according to the target rendering device.

Figure 4, demonstrates the current workflow of most graphical visualization applications and shows how VisArray attempts to simplify the workflow by abstracting commonalities.

5.2 Mapping Algorithm

The following content appears in a provisional patent application by the authors (Sweet and Sweet 2012). To map the point cloud data and image data together we need to first create a surface from the point cloud data. Analysis of the point cloud data from Cyclone, in the PTX file format, showed that the data is arranged as points in a vertical line with any cut out points being replaced with zero vectors. From this, we created an algorithm that looks at four points adjacent to each other and tests if there are any missing points. We then create triangles from all the points that are available. This algorithm is described in Listing 1.

Once we have the surface calculated, we can then progress to map the image data onto the surface, based on Figure 5.

We now discuss mapping 2D images onto the surface.

- **Determine two coincident points** A and B in the image and 3D data, we will assume point A is close to the center of the image. In Figure 5 the ‘noses’ of the two statues are used.

- **Find the viewpoint**. Find the vector from the camera origin O to point A, denoted vector C. Without loss of generality (w.l.o.g.) we assume here that point O is at the system origin

\[ O = (0,0,0). \tag{1} \]

- **Find the viewing plane**. Construct a plane P with normal vector C, w.l.o.g, we construct the plane at a distance of 1 from the camera origin O. Hence, using Eqn. (1), a point on the plane is \( \hat{A} = A/||A|| \) and the vector from the origin O to point \( \hat{A} \), \( \hat{A} - O \), is a unit normal vector to the plane. Then for point p on plane P, using Eqn. (1),

\[ p.(\hat{A} - O) - \hat{A}.(\hat{A} - O) = p.\hat{A} - 1 = 0. \tag{2} \]

- **Find the points in the viewing plane**. For each point \( P \) on the 3D surface, draw a line \( L \) from the camera origin O to the point \( P \) and find the point of intersection with plane \( P \), \( PP_i \). (Wolfram Mathworld 2010). For point l on line \( L \), and parameter t we have

\[ l = O + t(P_i - O), \tag{3} \]

which intersects the plane when
\begin{align}
t &= -1 - \frac{O \cdot (\hat{A} - O)}{(P_i - O) \cdot (\hat{A} - O)}, \quad (4)
\end{align}

Using Eqns. (1 - 4) this reduces to
\begin{align}
PP_i &= P_i / (P_i \cdot \hat{A}). \quad (5)
\end{align}

- **Find Y axis in the viewing plane.** Define a ‘vertical’ plane VP containing vector C, find the line Y where this plane intersects plane P. This will be the 2D ‘y’ axis. This can be accomplished by adding a vertical offset VOFFS to point A to give point \( V = A + VOFFS \), a new point on the plane \( \vec{V} \) can be calculated using the technique in Item 4) Eqn. (5)
\begin{align}
\vec{V} &= V / (V \cdot \hat{A}), \quad (6)
\end{align}

We require a unit vector \( \hat{Y} \) in the ‘y’ direction from the plane origin \( \hat{A} \) hence
\begin{align}
\hat{Y} &= \vec{V} - \hat{A} / ||\vec{V} - \hat{A}||. \quad (7)
\end{align}

- **Find X axis in the viewing plane.** Rotate the line Y through 90 degrees clockwise in plane P around the point where vector C intersects plane P. This will be the ‘x’ axis, which we denote line X. Given
\begin{align}
O &= (a, b, c), \quad V = (x, y, z), \quad (8)
\end{align}

then we find point H rotated around \( \hat{A} \) in plane P (Murray 2012) (for this method \( u = a, v = b, w = c \))
\begin{align}
dotp &= u \ast x + v \ast y + w \ast z,
H &= (a \ast (v \ast v + w \ast w) + u \ast (-b \ast v - c \ast w + dotp)
+ (-c \ast v + b \ast w - w \ast y + v \ast z),
b \ast (u \ast u + w \ast w) + v \ast (-a \ast u - c \ast w + dotp)
+ (c \ast u - a \ast w + w \ast x - u \ast z),
c \ast (u \ast u + v \ast v) + w \ast (-a \ast u - b \ast v + dotp)
+ (-b \ast u + a \ast v - v \ast x + u \ast y)).
\end{align}

We require a unit vector \( \hat{X} \) in the ‘x’ direction from the origin \( \hat{A} \) hence
\begin{align}
\hat{X} &= H - \hat{A} / ||H - \hat{A}||. \quad (9)
\end{align}

- **Project planar points onto ‘x-y’ axes.** Project each of the points PP onto the x and y axes to determine their ‘x-y’ coordinates PPX and PPY
\begin{align}
PPX_i &= (P_i - \hat{A}) \cdot \hat{X}, \quad PPY_i = (P_i - \hat{A}) \cdot \hat{Y}. \quad (10)
\end{align}

- **Find scale factors and offsets.** Calculate a scale factor S and ‘x-y’ offsets XOFF and YOFF for the 2D picture. Given that points A and B on the 3D point-cloud correspond to points \( \hat{A} \) and \( \hat{B} = B/||B|| \) in plane P, and assuming these points on the 2D image have ‘x-y’ coordinates \( (a_x, a_y) \) and \( (b_x, b_y) \) respectively, we have
\begin{align}
S &= ||B - A|| / \text{SQRT}((b_x - a_x)(b_x - a_x) + (b_y - a_y)(b_y - a_y)), \quad (11)
XOFF &= -a_x, \quad (12)
YOFF &= -a_y. \quad (13)
\end{align}

- **Calculate actual ‘x-y’ coordinates (pxy) for all points**
\begin{align}
px = PPX \ast S - XOFF, \quad ppy = PPY \ast S - YOFF. \quad (14)
\end{align}

Once we have these actual ‘x-y’ coordinates, we then have all of the information that we require to render a textured surface to the screen.

### 5.3 Examples

**Bond Hall Statue**

A Leica (Leica Geosystems 2010) scan of a statue at Bond Hall at the University of Notre Dame has been textured with a photograph of the same statue, taken from the same point as the scanner. The surface generated from the point-cloud data and the resulting textured surface can be found in Figure 5.

**Arch of Septimius Severus**

A Leica (Leica Geosystems 2010) scan of the Arch of Septimius Severus at the Roman Forum has been textured with multiple scanner
photographs from a number of scanner locations. The partially textured surface can be seen in Figure 7.

4. Method of Data Processing

In this section, we will describe the workflow required to transform a Leica scan of an object along with corresponding pictures, all the way through to the creation of the final viewable model. We assume that the reader has already registered their data within Leica’s Cyclone software (Leica Geosystems 2011).

The first stage is to extract the required data from Cyclone so that we can use it in our software. To do this, the registered point cloud data needs to be exported in the PTX format. This is due to the fact that other formats do not keep a rigid structure that allows the point data to be converted into a surface for applying the texturing. The next stage is to export every picture and the corresponding picture data from each “Scan World”, where each “Scan World’s” files are stored in separate folders.

Once we have extracted the data from Cyclone, we must first guarantee that the data that has been exported is correct. We run a utility on the PTX file to make sure that each model has the number of points specified at the start of the model description (the number of horizontal and vertical points that it should contain). This has become necessary since we have discovered that some large registrations have a small number of models that do not have the correct number of points compared to what we expect.

Once we have verified the data is correct, we then move on to creating a structure that our software can understand. This allows for the automated mapping of the pictures and surface generation. The first step is to convert the PTX file to a binary format. This reduces the data size of the file and allows for faster processing within the application. We then need to execute a utility in each of the “Scan World’s” picture directories to extract the required information from the photo information taken from Cyclone. This creates a corresponding info file for each image. Finally, we need to create a map file for each “Scan World” which contains the starting and end number for the pictures, which “Scan
Next, we can use our software to do the final processing. Firstly, the binary file is loaded into the software, which allows for the creation of the surface of the object. Next, we load each of the individual map files, which calculate which parts of the surface relate to which picture. Finally, the software writes out a valid DHZ (see Appendix A) file that can be loaded with the viewing application for visualization.

5. Conclusions

Mapping of a 2D image on 3D scan data is not a new discovery. It has been attempted by many in the last decade, with varying results. Our work is primarily intended to create an automated algorithm that combines the act of 3D scanning and photography, merging of the two in such a way that on site efforts can directly allow for 3D visualizations for scientists and research. This has many other areas of interest in the form of interactive education for high school students, who may not necessarily be able to visit sites such as the Roman Forum, and for students of architecture, history and archaeology that study such sites in greater detail.

We have evolved a workflow that optimizes on-site data acquisition and provides coverage by both data collection methods for optimal model/artifact coverage. The current techniques have been applied to test cases from the Roman Forum and Notre Dame as a proof of concept with software developed using the VisArray [http://www.visarray.com] framework for both a Java viewer and iPhone/iPad.

Future work will include extension of the DHARMA web portal [http://www.dharma3d.org] to allow researchers and students to upload their scanner data and photographs and access the combined model in the DHZ file format described in Appendix A. We also intend to submit the DHZ file viewer app to the Apple App. store as a means of viewing data in addition to the VisArray Java viewer.

Acknowledgements

Roman Forum Project: Office of Research, Office of Undergraduate research, Office of International studies, Center for Research Computing and School of Architecture University of Notre Dame funded this project. Dr. Fortini, from Soprintendenza Speciale per i Beni Archeologici di Roma, Ministry of Heritage and Culture and the Archaeological Service, Rome, Italy assisted the project.


Visarray team: Dr. Christopher Sweet, James Sweet and Dr. Kristina Furse.

References


Appendix A: File Formats

The authors have developed a file interchange format that optimizes the transfer of 3D data for both mobile devices and web applications. Because the proposed DHZ format is based on the proposed W3C web standard HTML 5 [http://www.w3.org/TR/2011/WD-html5-20110525/] markup language, micro-format and RDFa http://www.w3.org/TR/rdfa-syntax/ markup may be embedded inside the standard markup [http://dev.w3.org/html5/rdfa/]. These data attributes provide semantic information and vocabularies that allow machine agents to enrich the user experience by aiding in the exploration of the architectural models. By including references to linked-open data sources, users can discover background information such as place, history and architectural significance [http://www.w3.org/DesignIssues/LinkedData]. Because this data is linked much in the same way web documents are linked, the web becomes a query-able entity much like a standard relational database. For example, by using the Friend of a Friend (FOAF) [http://xmlns.com/foaf/spec/] vocabulary one can specify who was responsible for the capture and creation of a model. Geographic location can specified using Open Geospatial Consortium (OGC) [http://www.opengeospatial.org/standards/geosparql] standards allowing the use of web resources to query about place relationships. Lastly, links to architecture vocabulary can describe the underlying architectural elements, style and significance of the objects contained in the DHZ model.

A.1. DHZ File

The DHZ file format is a compressed container file that contains everything that is needed to render a processed model. The container is laid out with an XML file in the root that describes the model and its files. There is a directory that then contains the actual model data files, these include any number of mesh files as described in Section A.3, any number of point cloud files as described in Section A.4, and a number of image files if the mesh is textured.

A.2. XML Specification

To describe the models and their layout on the file system, we created an XML file format. The format contains models, containers, meshes and clouds. Models contain a title, a radius, a default camera view, the center of the model and a path to where the data files are stored. Models then contain a number of containers, meshes and clouds. Containers are a structure that allows grouping of clouds and meshes that all share a common transformation matrix and point size. Meshes contain the number
Listing 2. XML Example File.
of vertices, the number of indices, the name of a texture file to apply to the mesh, the file containing the mesh data, the file containing the index data, and a transformation matrix if it is not part of a container. Finally, clouds contain a number of vertices, the file containing the vertex data, and if it is not part of a container it will contain a point scale and a transformation matrix.

A.3. Mesh Data

Mesh files are stored in two files, both with the extension “.dat”, that are structured in a binary format. The first file is the actual mesh information that is laid out as described in Table 3. The second file is a list of indices into the mesh information and is laid out as described in Table 4.

<table>
<thead>
<tr>
<th>Position</th>
<th>Normal</th>
<th>Texture Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>Z</td>
</tr>
</tbody>
</table>

*Table 3. Mesh file format byte layout.*

<table>
<thead>
<tr>
<th>Polygon</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
</tbody>
</table>

*Table 4. Mesh index file format byte layout.*

A.4. Point Cloud Data

Point cloud files are stored with the extension “.dat” and are structured in a binary format with the layout described in Table A5.

<table>
<thead>
<tr>
<th>Position</th>
<th>Colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
</tr>
</tbody>
</table>

*Table 5. Point cloud file format byte layout.*
Investigating the Effectiveness of Problem-Based Learning in 3D Virtual Worlds.  
A Preliminary Report on the Digital Hadrian’s Villa Project

Lee Taylor-Nelms  
Booz Allen Hamilton, USA

Lynne A. Kvapil  
Butler University, USA

John Fillwalk  
Ball State University, USA

Bernard Frischer  
Indiana University, USA

Abstract:  
This paper discusses a recent study to test the effectiveness of combining 3D virtual worlds (VWs) with Problem Based Learning (PBL) in archaeological education of undergraduate college students at two American universities. The testbed used was a virtual world of Hadrian’s Villa at Tivoli (Italy), a World Heritage Site dating to the reign of Hadrian (117-138 CE). At both universities courses were offered on the villa using a PBL approach in such a way that the relative strengths and weaknesses of learning based on face-to-face, 2D, and VW presentations could be assessed. The study helped to clarify ways in which VWs can most appropriately be used as an aid to PBL.

Keywords:  
Problem Based Learning (PBL), virtual worlds, Hadrian’s Villa

1. Introduction

Hadrian’s Villa is the best known and best preserved of the imperial villas built in the hinterland of Rome by emperors such as Nero, Domitian, and Trajan during the first and second centuries CE. A World Heritage site, Hadrian’s Villa covers at least 120 hectares and consists of ca. 30 major building complexes (for a plan, see figure 1). Hadrian built this government retreat about 20 miles east of Rome between 117, when he became emperor, and 138 CE, the year he died. The site has been explored since the 15th century and in recent decades has been the object of intense study, excavation, and conservation (for a survey of recent work, see Mari 2010).

From 2006 to 2011, with the generous support of the National Science Foundation (NSF grant # IIS-1018512) and a private sponsor, the Virtual World Heritage Laboratory created a 3D restoration model of the entire site authored in 3DS Max. From January to April 2012, Ball State University’s Institute for Digital Intermedia Arts (IDIA Lab) converted the 3D model to Unity 3, a virtual world (VW) platform (Bartle 2003), so that the virtual villa could be explored interactively, be populated by avatars of members of the imperial court, and could be published on the Internet along with a related 2D website that presents the documentation undergirding the 3D model. The 3D restoration model and related VW were made in close collaboration with many of the scholars who have written the most recent studies on the villa. Our goal was to ensure that all the reconstructed elements—from terrain, gardens, and buildings to furnishings and avatars—were evidence-based.

Corresponding author: Tsn122@gmail.com
Once finished, the VW was used in two research projects. The first project involved use of the VW for some new archaeoastronomical studies. The VW was also used as a learning resource in two undergraduate classes offered at the University of Virginia and Xavier University. This article gives the results of the formative assessment of the effectiveness of the VW to further student learning.

Three-dimensional VWs present unique learning opportunities for students and instructors alike. Inside the VW environment students can operate an avatar in a historically accurate 3D space. The study sought to determine whether this new approach to learning was more or less effective than traditional resources (printed books, lectures, etc.) for understanding, communicating, and retaining information about complex historic sites such as Hadrian’s Villa.

Specifically, the study sought to measure the effectiveness of problem-based learning (PBL) in different types of learning spaces including the new one of VWs. As students presented solutions to problem-based learning scenarios in face-to-face, 2D web, and VWs, researchers compared student and instructor perceptions of communication, knowledge ownership, and problem-solving abilities. In what way does a learning approach affect the problem-solving, knowledge retention, and communication skills of study participants?

Does developing a PBL solution inside a VW deepen students’ communication, knowledge retention, and problem-solving skills?

2. Problem-Based Learning

Problem-based learning (PBL), which originated as a strategy for training physicians in medical education (Savin-Badin and Major 2004), is a teaching and learning strategy that is problem-centered. Students in classes in which PBL is used are challenged to develop answers to ill-defined problems with many possible solutions. Instructors facilitate the process of problem-solving. Rather than presenting facts in a lecture-style class, for example, PBL facilitators guide students in the problem-solving process, answer questions, point out helpful resources, and encourage discussion and reflection (Hmelo-Silver 2004; Hmelo-Silver and Barrows 2006). PBL is relatively new to archaeological education. The present study extends the previous work done by Kvpil 2009 while also applying VW technology to PBL for the first time.

2.1 Integrating PBL with VWs

The study was carried out in classroom-based courses at the University of Virginia in Charlottesville, VA and at Xavier University in Cincinnati, OH. Of the 37 students who agreed to participate in the study, all but one were full-time undergraduates. The 29 students from Xavier University were primarily freshman (17), with several sophomores (5) and juniors (5); there was one senior. The eight Virginia students ranged from freshmen to first-year graduate students. One student was a full-time graduate student in a teacher preparatory program. Altogether the study included 23 male and 14 female students; 36 students were Caucasian, and one student was African-American. All students were in their late teens to early twenties; there were no non-traditional students.
The course at Virginia was an art history seminar aimed at honors students. Achieving the learning goals of the course required the students to deepen their understanding of the art and architecture of the imperial Roman period. The class met once a week for a total of 2.5 hours. Because the class was an advanced seminar, students anticipated that independent research would be required. The course at Xavier was a Roman history and civilization course that focused on the imperial Roman period. The learning goals of the course required students to be familiar with the overall history, literature, and civilization of this period. The class met twice a week for a total of 2.5 hours. Because the course was the second in a sequence, the students expected that they would have to read ancient source material and conduct some independent research.

Since both courses dealt with similar subject matter, but originated from different departments and were aimed at a different kinds of students, the PBL component of the study was designed to find common ground between the needs of the two courses so that the appropriate content could be taught and the research goals of the study achieved.

2.2 Structuring PBL in the Syllabus

At the beginning of the semester, students were assigned to a team. Each week, teams researched and presented solutions to problem-based learning activities. Three of the nine solutions to PBL activities were presented in the face-to-face classroom; three were presented using a 2D web tool; and three were presented in the VW of Hadrian’s villa. Specific details on each assignment were released in class and each team was given exactly one week to research and complete an assignment. Face-to-face PBL assignments required students to present their solutions using only the affordances of the face-to-face classroom environment. 2D web problems encouraged students to use a range of electronic presentation tools such as Powerpoint, Prezi, Slideshare, or Keynote to communicate their solutions. VW problems allowed students to present their solutions as avatars of their choosing inside the VW of Hadrian’s Villa.

2.3 The Role of the PBL Facilitator

Studies have shown that one of the key success factors in implementing problem-based learning activities lies in the facilitator. A facilitator must be viewed by students as genuine, empathetic, and open to students’ intellectual work (Rogers 1969). A facilitator must be able to guide students in working together in groups to arrive at their own answers to complex problems (Rogers 1983). Finally, a facilitator must be able to foster a safe learning environment, motivating students to continue to take intellectual risks in front of an audience (Heron 1989). We endeavored to minimize the differences caused by various PBL teaching approaches by having the same teacher present each PBL activites to students in both classes. The role of facilitator in both the UVa and Xavier courses was filled by Dr. Lynne Kvapil, a scholar who had prior experience teaching PBL-centered courses and who was also the scheduled instructor of the Roman history and civilization course at Xavier. Dr. Bernard Frischer, a professor of Classics and Art History at Virginia and principal investigator of the research study, co-taught the course offered at the University of Virginia.

2.4 Problem-based Learning Materials

The students of both classes had access to a rich cache of resources for researching solutions to the PBL problems. These were made available on the Blackboard course site and through online resources including JSTOR and ARTstor. In addition, a number of print and digital resources were made available to students. Books about Hadrian’s Villa as well as Roman art, architecture, and history were put on reserve at the UVa Fine Arts Library and at the University of Cincinnati John Miller Burnam Classical Library.
3. Technology

The process of simulating the villa site of Hadrian was a significant interdisciplinary effort involving an international team of scholars, educators, designers, artists and technicians. Ball State University’s Institute for Digital Intermedia Arts (IDIA Lab) produced the multi-user avatar based VW of the villa employing custom middleware that leveraged the game engine of Unity 3 (figure 2). A primary goal of the project was to create environments that were both accessible and straightforward enough to engage first time users of VWs. The immersive learning environment could be experienced by using either installed executable versions and or from standard web browsers running on Windows or Macintosh operating systems. The back end administrative management relied on Smart Fox Server for creating accounts and regulating the 3D environment. The immersive environment was created through the integration and deployment of commercial products and custom software developed specifically for the project. The feature set available to students consisted of typed chat, a dynamic map displaying current visitor locations, online user status, teleportation via a map interface, multi-user voice channels, a triggerable avatar gesture system, layers of paradata, plan views, panoramas and site photographs, and integrated in-world URL links.

Visitors to the simulation entered the world by means of a custom avatar system designed specifically for the project. A user could select from a variety of avatars representing class, gender and ethnicity including the imperial court, senators, scholars, freemen, soldiers, and slaves. The avatar system was based on scholarly studies of circulation and flow throughout the villa. Avatars also served to populate the various places in the villa used for the daily activities of the court, including imperial audiences, dining, bathing, and worship. A visitor’s choice of avatar informed his social standing within the role-play of the simulation. A custom gesture system was created through motion capture and hand animation to provide the student with a unique set of actions and gestural responses specific to his chosen avatar. Gestures included greeting, affection, and supplication that varied in implementation according to the particular rank and class of an avatar. Communication technologies included public typed chat, private typed instant messaging and live group or private voice channels.

Integral to the project was the development of a companion web site for the VW of Hadrian’s Villa that provided learners with visual assets and written information about the state of the villa site today (figure 3). The visual assets include site plans, photographs, elevations, 3D models and panoramas of the archaeological park. The web site also provides descriptions and function of each villa feature and makes available scholarly interviews with prominent villa scholars. The web site offers paradata—a concept introduced by the London Charter (www.londoncharter.org/glossary.html)—making transparent the scholarship and methodologies of the reconstructions (from terrain to buildings, furnishing, clothing, and social protocols, etc.). All assets and information can be accessed by students directly from either the virtual simulation or from the web site itself.

4. Data Collection and Analysis Method

A review of literature from medical schools and universities shows that a purely qualitative
or quantitative form of assessment may fail to capture the added benefits experienced by students in PBL classes. Given that PBL activities have been found to improve not only a student’s knowledge of content but also higher-order thinking skills—skills often considered more difficult to assess—we developed a mixed methodology for collecting instructor and student perceptions of known PBL outcomes (cf. Barr and Tagg 1995; Ewell 1997).

Different forms of individual and group assessment were used to learn more about both group skills and individual knowledge. In general, assessment served as a learning tool given that three types of assessment were levied (cf. Earl 2003): assessment of learning (i.e., summative traditional tests—learn from a grade); assessment as learning (i.e., formative blogs—learn while participating in blogging activity); and assessment for learning (i.e., peer assessment rubrics—learn through feedback).

5. Preliminary Study Results

5.1 Observations of PBLs in Learning Environments

Each PBL activity required the students to solve a problem relating to the emperor Hadrian and life at his villa. For example, in the activity titled “Heads and Tails” students imagined they were officials in Hadrian’s governmental department that produced and distributed imperial coins. As members of this department, they were asked to design a coin that would promote him as emperor and reflect the values that inspired his construction of the villa. Students’ solutions to each PBL activity were presented in class to their peers and to the instructors one week after the activity was assigned. Each week, the instructors assessed student presentations and recorded observations. Assessment of the presentations focused on how students took advantage of each learning environment; whether students considered the impact and feasibility of their solutions; the perception of time, space, and structure in relation to the villa, and how well students could integrate prior knowledge with new knowledge in their presentations.

Changes in the way students considered time, space, and structure in the villa were most noticeable after students had been introduced to the VW because they had experience moving their avatars around the virtual space, and they had access to a date and time slider. Prior to that time, students were observed discussing buildings and their spatial relationships based on a traditional map of the villa. After students had access to the VW, they discussed not only buildings but movement through the spaces between buildings, and they gave their classmates directions based on surrounding structures and visual cues including corridors, staircases, statues, fountains, and other features that are difficult to visualize when reading from an architectural plan.

5.2 Formative Assessment/Blogs

Throughout the semester, students posted answers to formative assessment questions revolving around themes of communication, knowledge retention, and problem-solving skills. When students were asked what surprised them the most about the VW, one blogged: “while in the VW when working on
the PSAs with Maria (name changed), I am astounded by the graphics and the level of detail for each building. The VW has really helped me to understand the layout of the Hadrian’s Villa and what the buildings actually looked like. I feel like I have also finally gotten a grasp, especially after this specific problem-solving activity, on how to navigate Hadrian’s Villa and determine what routes you can actually take. The level of detail of the characters as well as the surrounding environment is also incredible. I also like how you can change the date and time settings, which is useful when picking a specific day for our PSA solutions. Ideally, more statues and art will be added” (Blog response, April 23, 2012). Another student blogged, “I find it amazing that the 3D world is shaped upon archaeologically supported details, and it makes me feel as though I am navigating in the past” (Blog response, April 30, 2012).

5.3 Focus Group Interviews

At the end of the term, student groups participated in topic-based, focus group interviews. A total of nine focus group interviews were conducted at both universities by educational researcher Dr. Lee Taylor-Nelms. Throughout the interviews, students stated that the concept of PBL and group work, in general, was not often used in other classes they had taken. They were more familiar with the traditional teaching construct of lecture, tests, and papers. The very idea that a PBL environment may not result in a single “right” answer proved unsettling for some. Several groups mentioned that the workload for this class was higher than that they had experienced in other courses.

When asked to compare differences in 2D, VW, and face-to-face PBL activities, students often commented on the spatial benefits of the VW. Although all students received a printed map of Hadrian’s villa, the ability to navigate inside the reconstructed buildings in the VW gave them a different perspective on how historical events might occur. In addition, since the VW did not permit avatars to fly, walking around the villa as an avatar gave them a better appreciation of the scale and dimensions of the villa. “It took a long time to get anywhere,” one student commented (Student interview, May 1, 2012). The presence of non-player avatars (avatars not controlled by humans) also gave students a better sense of how the space worked. It made them ask questions that may not have occurred to them otherwise such as: “How many groundskeepers did it take to run Hadrian’s villa?” (Student interview, May 1, 2012). Several groups mentioned that they were able to control the positioning of the sun inside the VW space. This control allowed them to see how the space looked at specific points in time and proved beneficial in better understanding the interrelationship of culture, religion, and architecture.

Several groups noted the similarities in the face-to-face and VW PBL activities because required more group collaboration than did the 2D activities. Although instructors did not prescribe how to deliver presentations, students interpreted face-to-face and VW learning environments as exercises where role-playing was appropriate or even expected. Some groups mentioned that not all activities lent themselves easily to role-playing solutions. In those cases, students found a face-to-face or VW solution to be limiting. Other groups described role-playing solutions as enlightening because they offered a new perspective on a problem. “By re-enacting a solution, we could actually see what worked and what didn’t,” one student remarked (Student interview, April 30, 2012). Many noted the biggest difference between VW and face-to-face activities was the opportunity to bring physical objects into role-playing scenarios. In some face-to-face presentations, students dressed in traditional Roman garb or prepared food for a feast. Although it is technically feasible to dress an avatar or to arrange for a virtual feast inside the VW, creating objects and altering an avatar’s dress is not an available option in the current version of the VW of Hadrian’s villa.
Almost every group mentioned that face-to-face and VW PBL solutions took longer to prepare and present than did the 2D presentations. The reason was that groups felt they needed to coordinate and practice their “parts” with other group members. In contrast, with 2D slide presentations, the groups felt they could simply break the presentation down into discrete sets of slides on the various topics to be covered, and give each student in the group the responsibility to create and present a given set of slides. Presentation time of VW and face-to-face activities often took longer because students felt 2D slide shows allowed them to use bulleted points or images to express multiple solutions to problem, whereas VW and face-to-face solutions required explicit, verbal expression.

5.4 End of Semester Instructor/Student Surveys

End-of-term instructor surveys saw student success as based on understanding ideas more than memorizing facts. In addition, instructors thought students found the VW engaging and motivated them to spend more time exploring the VW and the resources in it. Instructors agreed that learning how to use the new VW resource and dealing with inevitable technical difficulties diverted some attention from academic tasks. However, by the end of the course, they believed that their observations of student PBL solutions in VWs attested new ways in which students thought differently about issues such as space, time, and structure.

End-of-semester student surveys asked students to respond to similar questions posed to instructors. However, it was clear that student perceptions on the effectiveness of the VW environment was more influenced by some of the technical difficulties many experienced in using the VW. Despite the difficulties, both instructors and students “extremely agreed” that the scholarly team should continue to keep improving and extending the Unity 3 application. One student wrote: “the 3D world has potential and it can grow into something if it is explored and more options are given to students who actually desire to make it work. All ideas have to be listened to and explored. It requires a teacher and a student to experiment with it and find out new creative ways to use it outside of its intentions.”

6. Conclusions

PBL activities presented new challenges for many students at the two universities. For most students, the learning environments in which PBL activities were presented proved less challenging than the concept of PBL itself. The very notion that a problem could result in multiple solutions and that students could learn from each other struck some students as downright antithetical to pedagogy in the humanities. Despite the fact that PBL has long been in use in other fields, PBL struck many students (especially those not majoring in Education) as a new and experimental approach to learning. When asked, students often reported that with the exception of science labs, most of their college classes did not require group work and classroom presentations. Their typical classroom experience had consisted of lectures, tests, papers, and quizzes.

Thus, it is not surprising that when asked to be creative and take advantage of the unique affordances of each learning space, most students lectured. With encouragement from the PBL facilitator, groups began to role-play and write skits rather than to deliver reports. Some groups even dressed up and brought in props like food to make their role-play more authentic and meaningful, but most did not question the primary delivery mode which was (in more cases than not) an oral delivery. Students could have used space as a catapult to present information in creative formats but often they did not choose to do so, in part because of the limitations presented by the infrastructure of the classroom. In one face-to-face presentation, a group used the desks in the
classroom to simulate the layout of a Roman banquet, but more frequently students seemed to feel constrained by the traditional classroom layout.

2D presentations were similar in that students reported what they knew with the help of visuals. Being able to use Powerpoint or Prezi as a visual organizing tool seemed to help some groups to communicate information more effectively. In most cases, students noted 2D presentations were easier to follow because the information was clearer and easier to understand. It also seemed to help groups in managing work flow virtually without the added pressure of finding time for face-to-face meetings. Also, 2D presentations seemed to offer a familiar, comfortable format for most students. In effect, when asked to give a 2D presentation, students felt they were playing the role usually filled by their professors.

VW proved the challenging for some students because it required them to be creative in ways traditional classroom formats do not. Given that the VW resource was designed to function best as a tool for spatial navigation, it greatly facilitated archaeological problem-solving, allowing students to develop a spatial understanding and appreciation of problems set in ancient times. As beneficial and unique as this benefit is, this single feature did not always result in superior communication or knowledge retention than did the more familiar formats of 2D slide shows or face-to-face presentations. To some extent this may be attributed to the trepidation students felt toward unfamiliar technology. More experience with VW technology and role-playing via avatars might encourage more creative solutions while giving students the opportunity to have the experience of operating in an ancient environment.

In terms of formative evaluation of the digital resources created to support the educational research, both students and instructors agreed that the VW needs additional improvements to become fully functional in promoting the PBL goals of communication, problem-solving, and knowledge retention. Specifically, students felt the following features would greatly enhance the value of the VW for learning:

- Enable students to switch between avatars without having to open/re-start the program.
- Add an orientation map in the villa so it is clear where you are at all times.
- Add a sound signal over avatars so it is clear who is talking when.
- Add the ability to switch the view of an avatar from 1st to third person.
- Add the ability to zoom out from your avatar and focus in on other aspects of the world.
- Put the avatar name above the avatar so others know who is who.
- Allow students to click and receive information on avatars (bio, clothes, animations).
- Permit multiple avatars (controlled by students) in-world without technical problems.
- Allow objects to be picked up and include metadata about those objects.
- Develop realistic animations with non-player characters (NPC), e.g., avatars stab a NPC, who reacts.
- Include information about NPCs when clicked.
- Allow avatars to move faster i.e., to jog, run, or walk faster.
- Allow avatars to build objects, e.g., to create a banquet.
• Permit customization of avatars i.e., dress, animations, and gestures.

The IDIA Lab will be taking advantage of the testing of the VW reported here in order to make improvements for the next release, which is scheduled for the fourth quarter of 2012. In fall 2012 and spring 2013, this version will be tested in courses taught at the University of Virginia and Xavier University.

The main lesson learned thus far is to confirm that PBL can be effectively used in archaeological education, albeit in a new way. In the past, PBL has been used to put students in the role of the contemporary archaeologist researching ancient material (Kvapil 2009). In the present study, PBL promoted learning by putting students into the role of ancient people.

At this time, the communication capabilities of the Hadrian’s Villa VW are limited compared to a face-to-face presentation environment in which students can convey information non-verbally by displaying objects, changing one’s dress, and re-inventing a scene. The VW, as currently designed, is also fixed and limited to oral expression through voice or text. From the learner’s perspective, solutions presented in the VW may have been more difficult to follow given that solutions could only be expressed within a pre-staged environment without alterations. Learners found a 2D presentation easier to understand as it allowed students to share information visually (through pictures/symbols), and reinforced key oral information in bulleted points, but this viewpoint also reflects the expectations of students who had much greater familiarity with traditional class structure and the technology that accompanies most art and archaeology courses. Students found 2D and face-to-face environments more effective in communicating and hence, understanding and retaining knowledge in the long-term. But the richness of the VW environment of Hadrian’s Villa provides a new opportunity for students to explore and master complex Roman architectural spaces and to participate first-hand in the functioning of ancient society in ways that cannot be replicated even if each student had the opportunity to visit ruins of Hadrian’s Villa.

Acknowledgements

We thank all the many scholars who kindly contributed to the research on the 3D reconstruction model, on the VW of Hadrian’s Villa, and on educational technology generally, including: Dean Abernathy, Benedetta Adembri, Mary T. Boatwright, Federica Chiappetta, Anthony Corbeill, David Dearborn, Marina De Franceschini, Stephen Ehrmann, Fulvio Cairoli Giuliani, Robert Hannah, Jens Koehler, Paolo Liverani, Zaccaria Mari, Elizabeth Macaulay-Lewis, Robert Mangurian, Zaccaria Mari, Mary Ann Ray, Stephan T.A.M. Mols, Daniel Wagner, Leslie Yarmo, and Michael Ytterberg. We single out for special commendation Matthew Brennan for his fine work in overseeing creation of the 3D reconstruction model of the villa. We thank the Soprintendenza per i Beni Archeologici del Lazio and Superintendent Marina Sapelli Ragni for permission to gather and publish documentation on the site of the villa. Finally, we are grateful to a donor, who wishes to remain anonymous, for making it possible to start the project and to the National Science Foundation whose support (grant # IIS-1018512) made it possible to bring it to a successful conclusion.

References


Building Blocks of the Lost Past: Game Engines and Inaccessible Archaeological Sites

Anna Maria Kotarba-Morley
University of Oxford, UK

Joe Sarsfield, Joe Hastings, John Bradshaw
Nottingham Trent University and Angry Pirate Productions, UK

Peter Nicholas Fiske
University of Oxford, UK

Abstract:
This paper explores an idea for creating an informal and easily approachable media platform to promote archaeological sites that are inaccessible and lesser known to the public in the form of an educational game. This game will create an illusion of a real archaeological site visit, allowing players direct contact with its environment and surroundings as well as interaction with its ancient and contemporary inhabitants. In an era of international connectivity, globalization, and social networking, it seems appropriate to choose the online computer and mobile gaming industries as media for spreading the interest in heritage and archaeology.

Keywords:
Heritage Education, Public Archaeology, Computer Game, Inaccessible Archaeological Sites, Media in Archaeology

1. Introduction

Games lubricate the body and the mind.

—Benjamin Franklin

The aim of this paper is to explore whether an online educational computer game based on scholarly, detailed, and accurate 3D reconstructions of archaeological sites is a viable medium for displaying those sites which are inaccessible to the public. To this end, this paper will initially focus on reviewing the market for computer games and commenting on current games which are based around historical and archaeological environments. It will continue by discussing the platforms (e.g. PC, console, tablet, smartphone, museum stand) that may be used, whilst assessing their capacities and their potentials for application to different projects. Next, it will discuss the merits of existing game engines for the creation of such games and present frameworks for games which may be applied to different target groups, client needs, and scenarios. It will additionally seek to answer how to provide successful ways of developing games as tools for education through entertainment.

The idea for creating an interactive medium to explore archaeological sites developed from one author’s work on a site in an area of the Middle East which is inaccessible to the general public. There, the local environmental conditions would pose additional cost and maintenance challenges for the open-air exposition of the site. Such a site would form a rewarding basis for a case study, although one has not been attempted here in favour of first formulating the principles and methods by which a study would be undertaken. Through discussions and comparative analyses with other sites, it became clear that there is a great need to investigate new media for the displaying and sharing of cultural heritage of all kinds, on the land and under the water (Kalay et al. 2007).
One of the key rationales for archaeological investigation is to preserve the legacy left to us by previous generations. Another is to make this heritage available to future generations. The computerized, globalized, and networked 21st century obliges us to explore new frontiers in archaeology for the public. Public interest in archaeology is clear and sustained. Film and video game series such as Indiana Jones and Tomb Raider have left indelible marks on popular culture, becoming key reference points for archaeology in society (Drew 2007).

Nevertheless, traditional displays like those in the older galleries of the British Museum still remain extremely popular, placing this institution at the top of the list of most-visited British tourist attractions every year. The most recently published BM annual review shows that it had 5.7 million visitors, almost 200,000 more than in the previous year (Rickman and Wilson 2010). Despite this influx of visitors, institutions like the British Museum are also seeking new, digital ways of displaying their collections online. For example, the BM’s public-access database maintains information on almost 2 million objects and has been visited by over 15 million people for general interest and research purposes. Those numbers show how popular online resources are becoming in teaching and exhibiting cultural heritage, confirming that the internet is rapidly becoming the medium of the future.

Museums today are investing in online games and multimedia displays, such as interactive stands, 3D reconstructions, and holograms of ancient buildings or archaeological objects. Internationally recognized sites, such as the Roman Baths Museum in Bath Spa and the archaeological site of Pompeii, might seem very easy to present to visitors. Their outstanding cultural values and easy site access have already been recognized by UNESCO with World Heritage Site status. Nevertheless, even at such phenomenal monuments, museum managers and exhibition designers are always in search of new methods of making sites more attractive to visitors. Lesser-known, inaccessible, and under-funded sites, however, do not get as much attention. Creating a game based on accurate reconstructions of past environments and peoples would be a rapid and cost-effective way of displaying such sites and making them available to broad groups of new visitors – players.

2. Gaming and Archaeology

There is a large body of games available on the market which relate directly or indirectly to archaeology. These include games which are set in or based on particular historical periods and rely on archaeological and historical information to create the game’s environment, such as Assassin’s Creed: Brotherhood and the Prince of Persia, God of War, Total War, and Age of Empires series. They also include games in which the player controls a modern character exploring an ancient environment, such as Tomb Raider or Unchartered.

Games set in ancient contexts are predominantly strategy games allowing the player, for example, to control an army equipped with contemporary weapons (e.g. Rome: Total War, in which the player controls legionnaires, or Age of Empires II, in which the player controls a medieval army). These games often inadvertently educate the players in military history, exposing them to landscapes and technologies that existed during particular periods. It is important to note that many of these games are by no means accurate and better serve to raise players’ interests in history rather than operate as educational tools per se. They also tend to support traditional historical narratives which emphasize the political and military aspects of the past.

Strategy games are most popular with PC gamers. In 2011, 33.6% of all PC games sold were strategy games, in contrast to just 3.8% strategy games sold for console (Gallagher 2011). Strategy games often require complex
controls, and for this reason are not as popular with the console market. Successful examples of these games include the *Total War* games (*Shogun, Napoleon, Empire, Rome*), *Rise and Fall: Civilizations at War*, and *Sid Meier's Civilization* (I–V).

In the series of strategy games called *Sid Meier's Civilization*, players must lead their civilization from antiquity into the future, researching different technologies along the way. In the game, players control armies from different cultural periods. The latest *Civilization* game in the series, *Sid Meier's Civilization V*, was the 5th best seller in 2010 on Steam, a popular online games store, with an estimated gross revenue of $21.9 million (Tassi 2011). It is also currently Amazon’s top-selling PC strategy game (followed by *Total War: Rome II*) (Amazon.com 2013), demonstrating that games with archaeological or historical themes are popular and are successful on today’s market.

Games set in modern contexts are predominantly adventure games in which the protagonist is an adventurer or archaeologist who is exploring ancient ruins, such as in *Tomb Raider: Unchartered 3*. Other successful examples of these games include the *Total War* series and *Rise and Fall: Civilizations at War*, which are based in several different eras, and games like *Spartan: Total Warrior*, *Dynasty Warriors*, *Prince of Persia*, and *God of War*, which are mainly based in one historical period or in a mythical past. These games are more popular with console players; 29.2% of all console games sold in 2011 were action or adventure games (Gallagher 2011).

Statistics show that the market for these games is extensive, with an expected total consumer spending in the games industry in 2012 of $25.1 billion (Gallagher 2011). The fact that 82% of 8 to 65 year olds in the UK and 68% of American households play computer games (GRABstats.com 2012) shows that the receiving audience is not only significant, but also very diverse. The numbers also show that the average age of a gamer is 30 and that 62% of gamers play with others, either in person or online (Gallagher 2011). These and other statistics also help to target games to particular types of people.

Educational archaeology games for kids are widely available on the market. Some of them are created by museums, such as games and educational media from Birmingham Museum and Art Gallery, *Dirt Detective* by the Museum of Williamsburg, *Londinium* by the Museum of London, *Every Coin Tells a Story* from the Cotswolds Museum, and *Roman Design*, *The Mosaic Game*, and *Across the Board* by the Tyne and Wear Museum. These online games are usually targeted at players aging from 5 to 15 years old and serve not only to educate, but also to develop general interest in the past and encourage direct interaction with archaeological material through museum and site visits.

Some other browser-based games run via services like the BBC, such as *Mummy Maker*, in which players learn how to prepare a mummy; *Dig Deeper Quiz*, an archaeological quiz; *Hunt the Ancestors*, in which players run archaeological prospection, excavations, and buy specialist reports whilst on a limited budget; *Diver's Quest*, in which players dive into the deep to reach a treasure chest whilst answering questions about underwater archaeology; and *Dig It Up: Romans*, in which players move three diggers equipped with a geophysical survey around a field in search of Roman artefacts. These games build archaeological and historical awareness in children and teenagers; however, they are constructed in a way that they can be played just once or twice because the sets of questions and tasks are very limited.

In recent years, there have been several successful recreations of past cultural heritage environments using different game engines (Champion 2011). These range from video fly-throughs of pre-rendered 3D landscapes to
interactive environments, or those accessible for example through the touch screen display (Wang and Champion 2011). Players can walk around and experience sites with their features and objects and immerse themselves in the past environment as if they were a part of it (Champion 2004). Additionally they have an opportunity to explore both, tangible and intangible past – through interaction with e.g. buildings, archaeological features, and landscapes; or through interactions with inhabitants of the past place, their rituals, dances, traditions (Kalay et al. 2007). The game, if successful, becomes ‘engaging, challenging, and intuitively usable, and helps create a fantasy world for participants’ (Champion 2006).

The *Second Life* engine has been used many times to recreate archaeological and historic sites. At the peak of its popularity, *Second Life* allowed for an exceptional level of interaction and movement through the environment, but the visuals are no longer up to the standards of other game engines. One of the most superb reconstructions in *Second Life* was of Çatalhöyük. Created by UC Berkeley, it won an Open Archaeology Prize (CIO 2007). However popular as it was for its time, the *Second Life* engine is no longer at the cutting edge of gaming technology and has fallen out of fashion. The features of the *Second Life* engine do not match the current generation of game engines. This is why the information for displaying particular sites should be created and curated in such a manner that the raw data can be used in different gaming environments.

Another type of visualization uses point cloud data, usually gathered from laser scanning with an optical remote sensing technology known as Lidar, for fly-throughs or 3D models. The immense level of detail makes it necessary for them to only be presented as pre-rendered fly-throughs, with no user interaction at all, such as the fantastic reconstruction of Musawwarat es-Sufra in Sudan (Rüther 2010). Whilst these are extremely interesting for specialists or amateurs fascinated by a particular site or period, they are not environments which invite visitors to explore them in detail and in multiple visits, unless they are conducting specific research. Nevertheless, they show a value and a potential to be transformed into environments which can be accessed and explored by the general public.

The Portus Project (www.portusproject.org) by the University of Southampton is a great example of a pre-rendered fly-through (Earl et al. 2008; 2009; 2010) showing an accurate, high-detail recreation of the great port of Imperial Rome. The Southampton team used a combination of Google SketchUp, Google Earth, Autodesk 3ds Max, and *Second Life* to create the environment. Project creators used a great mix of previous data from archaeological surveys and excavations, along with detailed library research, to create buildings and environments to scale. Although the environments created serve their purpose as accurate representations, if an environment is to be used for a game or any kind of interactive activities it requires additional details. Those details need to create an illusion of reality and bring a place to life with features such as trees, plants, dirt on the streets, and humans. Modern game engines have tools for creating content like this which can be used to add these details relatively easily. Such as Unreal Engine 3’s Instanced Foliage System (Epic Games, Inc. 2012), which allows the developer to select the tree or bush that they want and paint a forest of them straight onto the landscape, automatically rotating and varying the scale to give a realistic feel.

Underwater cultural heritage exhibits an underestimated and underexplored potential for gaming. Oceans and freshwater cover over 70% of the total surface of the earth and the estimates suggest that 3 million shipwrecks still lie on the seafloor. The 2001 UNESCO Convention for Underwater Cultural Heritage stresses that *in situ* preservation is ‘the first and preferred option before allowing or engaging in any activities directed at this heritage’. In light
of this recommendation, access to most of the underwater cultural remains will be very limited and usually possible just for visitors with diving certificates. This area of heritage studies is therefore very much in need of further digital and interactive development. As mentioned above, there are numerous games for children and teenagers about underwater heritage; however, high-end strategy or adventure games based on detailed and archaeologically accurate 3D visualizations are not yet available on the market.

In the light of the examples which have been presented, it is clear that a high-end educational adventure or strategy computer game based on accurate reconstructions of archaeological sites is needed. It also shows that it has a great potential in the market to compete with games based on inaccurate or fictional representations of the ancient world.

3. Proposals and Possibilities for an Educational Game Project

This section aims at proposing a framework for developing an educational computer game which could be designed to complement the 3D visualisations of archaeological sites and allow general public exploration of virtual domains. At a later stage, this method could be applied to the development and testing of a case study based on an archaeological site which is inaccessible to the public. The game can also be tailored to different audiences and the users of different devices, however it is envisaged that the target age group will be 15 to 45 years old and the target device will be the PC.

This kind of game should have no winners and each level could be just a subsequent chronological epoch. The reward could simply be to interact with the environment of the next period; this is also the case with a popular recent release, *Journey*. There have even been mods released for some games, such as *The Elder Scrolls: Morrowind*, *Oblivion*, and *The Elder Scrolls V: Skyrim*, which turn off encounters with enemies allowing the player to explore the environment freely with no interruptions. The player could collect and look at artefacts around the environment along the way, giving them a better understanding of the historical environment. Given the subject matter, which is dedicated at a particular target audience, it is less likely that the game will become a blockbuster. Nevertheless, it will allow for archaeological reports, which are sometimes difficult to comprehend for non-specialists, to find their way into the general knowledge.

3.1 Methodology

In the current climate it is possible to have the best of both worlds: there are free game engines available with powerful 3D rendering capabilities. Although they may still not match point cloud data in resolution, with efficient level design and content creation they can provide a very high level of detail. The added benefit is that these game engines can all run in real time, allowing players to interact with the environment as they please and explore it at their own pace (Calef et al. 2002).

3.1.1 Collecting Data Needed for Visualisations

Source data is needed before creation of the game’s environment can begin. Many archaeologists long ago saw the usefulness of presenting their sites and findings through 3D visualisations, and those will usually form a perfect environment for a computer game. Additional information can be gathered from photographs, videos, laser scans, point cloud data, excavation reports, maps, technical drawings, and existing historical and comparative information. These methods can also be used to create a 3D model from scratch.

Laser scanning is an extremely useful form of data collection, reproducing existing objects, buildings or entire landscapes in near-perfect detail. Data is collected in the form of a point cloud, a list of millions of points, which when
fed into a 3D development environment can produce a model of extremely high detail and accuracy. As these models have a huge amount of faces, or polygons, they are not suitable for use within a game engine, slowing it down and not allowing for performance in real time. An artist would have to take the scan and break it down into segments in order to recreate the model more efficiently. This process would take a considerable amount of time, but it would still be much faster than creating the same models using only photos or drawings.

Conversion from point cloud to a 3D model that can be used in a 3D development environment, for example Autodesk 3ds Max, consists of four stages: pre-processing, determination of surfaces, polygon generation, and post-processing (Remondino 2003). These actions can be undertaken by use of software such as Pointools or Autodesk AutoCAD. In order to then make the resulting models suitable for use in a game engine, a considerable amount of simplification would have to take place in the post-processing phase. The simplification process (Weyrich 2004) consists of removing a large amount of faces from the model whilst still keeping a desirable level of aesthetic quality. Having a smaller amount of faces to draw will allow computers to deliver the scene to the screen faster, therefore maintaining high performance (Arnaud 2011) and ultimately allowing the user to experience environments in real-time. Detail removed in the simplification process can be replaced by the creation and application of textures using techniques such as normal or displacement mapping. These techniques give the appearance of a more detailed object, whilst keeping performance costs significantly lower than using the original, highly detailed model.

Once the model has been completed, data from the original scan can be used to create high detail textures, giving the finished model a realistic appearance. This process is extremely useful for recreating sites or artefacts that are complete and undamaged, or where the desired outcome is a 3D representation of a ruin or a damaged object. It is less useful for recreating objects or features currently not available but known from the historical record.

It is important that as many pieces of source data as possible are collected when creating a 3D environment or representation. Relying on one source too heavily may give an inaccurate outcome. Even with a great deal of data available, it can be difficult to achieve a comprehensive understanding of how buildings or objects looked in their original conditions. A certain amount of artistic license may have to be used in order to give these environments a life-like feel if they are to be made into an interactive game environment.

3.1.2 Interactivity

The level of interactivity would vary according to the specific needs of the project. If it is a game destined to be used in a museum stand, then it will require a very different form of interaction from a project meant for a personal computer as a research tool, for education, or for entertainment (Amory et al. 1999). A game designed for a PC or a games console could either be interacted with by use of a keyboard and mouse, or a gamepad. Use of a gamepad, however, is preferred over the use of a keyboard and mouse combo, because it is more comfortable for the players.

A project destined for a museum stand would work better with a more immersive level of interaction. Features such as 3D and a motion sensor could be incorporated. The motion sensor could be used to control an on-screen character or navigate menus and interact with data such as graphs and movies. The players could also interact with specific virtual objects, e.g. a piece of pottery, which they can pick up, move around, and learn more about.

A game may also be solely meant for mobile devices such as tablets or smartphones. This could be useful as a data acquisition tool with which professional archaeologists
could acquire information on their mobile or tablet device when working on a site. Such an approach is already used in archaeology, such as at Pompeii. Drawing/CAD tools can be useful on these devices at archaeological sites and unlimited canvas size and huge zoom levels allow for great detail when creating scale drawings’ (Wallrodt and Tucker 2011; Tucker and Wallrodt 2013; Wallrodt et al. 2013). Data collected in this way could be transformed further to allow the user visiting a site to see how it looked in a particular period of the past. Tablets rigged with GPS can position the user in a particular point on the surface of the earth and via an online connection with the game engine to display what he sees in front of his eyes, but in the past. This approach has been successfully implemented in Gunnar Liestøl’s visualisations for iPad, which he presented at CAA 2012 (Liestøl 2009).

Another possible use for a visualisation on a mobile device could be an as educational game, which is likely to reach a wide audience. With the rise in popularity of smartphones, they have become favoured devices to play games on. The rise in power of smartphones in recent years has also allowed full 3D environments to run on them in real time. Although the performance does not yet match that of games consoles or PCs, a lot can still be achieved on a mobile or tablet device when taking into account the performance limitations during game development.

3.2 A Possible Game Engine

When it comes to creating a game using a 3D visualisation of an archaeological site, an engine is needed to combine all the models and add functionality. Modern game engines also allow for effects, such as real-time lighting. Post-processing adds effects after the game has been rendered, such as depth of field, colour correction, and motion blur. Particle effects simulate real-world effects, such as fire, flowing water, or smoke. Users may interact with the environment through keyboards, mouses, gamepads, motion capture devices such as the Kinect, or combinations thereof.

The three most suitable free and currently available engines are outlined below: Unreal Engine 3, Cryengine 3, and Unity (Table 1). New engines are released frequently by many companies, so only currently available ones with free SDKs (Software Development Kit) will be assessed. Deciding on the correct game engine is an important step in development as each engine has its own strengths and weakness.

Unreal Engine 3 would be the more appropriate engine to accomplish this project. The visual scripting language Kismet and state of the art graphics allow for rapid development of interactive environments and realistic representations of historical sites. The reasonable costing structures make the engine suitable for most financial situations. Although as of this writing Unreal Engine 4 has not yet been released, it is currently in production by Epic Games and a video demonstration of its capabilities was presented to the public in mid-2012. It remains to be seen precisely when the software will be made available and whether it will have a free SDK. Naturally, the latest version of the Unreal development environment available would merit consideration for this or other projects.

3.3 Possible Scenarios

There are many areas within archaeological and heritage projects and institutions where...
educational games can be used. The most obvious use of a game would be to recreate a historical site as it would have appeared in its own time. In such a game, the player would be able to navigate around an environment and interact with different areas, buildings, or objects in order to learn more about them.

The player would also be able to explore the fluctuating fortunes of the site, seeing it at war, at peace, in bloom, and in decline. This could be expanded upon by creating storylines, quests, or tasks for the user to play through. This will promote learning as the player would be encouraged to continue playing in order to complete the tasks.

A second scenario in which a heritage game would be useful would be to display a historical site as it appears now. This would be particularly beneficial in museum exhibitions and on-site displays. This would allow players to see and explore the place where the 3D artefacts within the game came from. The exact location and the context of the findings could be indicated to the users giving them an insight into the archaeology behind the exhibit and a greater understanding of the work that archaeologists do (Fig. 1).

A further use would be to create a library of historical objects in which players would be able to select an artefact and be shown relevant information and 3D reconstructions of the

---

**Table 1. A comparison of the three game engines under present consideration.**

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unreal Engine 3</strong></td>
<td></td>
<td>If a project was funded and needed to be produced quickly for multiple platforms, then UE3 would be the best choice. The editor for the engine allows the developer to quickly bring a concept together into a full game and build out for different devices. As long as performance limitations of mobile and tablet devices were taken into account from the start, few changes would have to be made to the product for a PC version or mobile version.</td>
</tr>
<tr>
<td>easy to create for multiple platforms at once</td>
<td>once $50,000 is earned in revenue from a product, 25% goes to Epic Games</td>
<td></td>
</tr>
<tr>
<td>supported on most modern smartphone and tablet devices</td>
<td>slightly outdated graphics</td>
<td></td>
</tr>
<tr>
<td>free to develop for PC + iOS up to $50,000 earnings</td>
<td>must pay $50,000 for full UE3 license to develop for Android devices</td>
<td></td>
</tr>
<tr>
<td>powerful and robust engine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>social integration using Game Center, posting to Facebook/Twitter walls</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cryengine 3</strong></td>
<td></td>
<td>If a project was purely intended as a highly accurate visual representation of an environment, then Cryengine 3 would be the best choice. This is because it has the best graphical rendering capabilities of the three game engines compared.</td>
</tr>
<tr>
<td>real time in-game editing</td>
<td>not supported on mobile or tablet devices</td>
<td></td>
</tr>
<tr>
<td>current gen graphics</td>
<td>must pay for full licence if end product is made to be profitable</td>
<td></td>
</tr>
<tr>
<td>free to develop for PC if end product is non-profit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>give 20% of revenue if end product is profitable</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unity</strong></td>
<td></td>
<td>If a project had a large time scale and a low budget, then Unity would probably be the best choice. This is because it is relatively cheap to develop for any platform to develop for any platform, in comparison to the other 2 options mentioned; however, it is not as developer friendly or powerful as UE3 or Cryengine 3.</td>
</tr>
<tr>
<td>relatively cheap to develop for PC, iOS and Android</td>
<td>requires 3rd party software for many features such as multiplayer or posting to Facebook and Twitter</td>
<td></td>
</tr>
<tr>
<td>supported on most modern smartphone and tablet devices</td>
<td>outdated graphical capabilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>more difficult to develop with</td>
<td></td>
</tr>
<tr>
<td></td>
<td>developers support for engine does not match UE3</td>
<td></td>
</tr>
</tbody>
</table>
object, much like the inventory in games such as *The Elder Scrolls V: Skyrim* (Fig. 2).

This view would allow users to see the object from all angles and interact with it in a way that would not be possible in a museum or heritage site. One approach would be to have a game with two storylines, one which could be set in antiquity and one in modern times.

The potential player could choose to play one of the available ancient characters, such as a warrior, a priest (Fig. 3), a sales-woman, an artist, a thief, or another created according to the particular specifications of the game. On the other hand, the player could also choose to become a contemporary character, such as a tourist or an archaeologist. This kind of game would be aimed at exploring sites in different periods and through the eyes of their different inhabitants, humanizing them and leading to a better understanding of history by the player. It would also allow for just a simple tour of the site for those players who are not normally interested in gaming. However, for those who always dreamt about being like Indiana Jones, there could be an option of becoming an archaeologist character and taking part in fictional excavations. Below, we present a simple demonstration of how the 3D environment set in an ancient Roman provincial port could look like (Fig. 4). This image is only meant to provide a sample of the technical possibilities; it is not meant to represent a historically accurate reconstruction reflecting any archaeological or scientific evidence.

It would also allow for direct interaction with other online players in the form of a massively multiplayer online role-playing game (MMORPG), all of whom continuously create, re-invent, and populate the site's universe. For example, when building a house, players would be able to acquire the materials with which to build it, source the labour and skill necessary for its construction, and buy or build the furniture to make it a home. The players would be allowed to set their own tasks and, along with their co-players, create and live in an ancient community. The game designers would make sure to pre-set original tasks in a way to keep the players motivated to continue and explore new levels. Some statistics report that the average player before the age of 21 has already spent 10,000 hours playing video games and mastering new tasks. However, schools and educational institutions often struggle with
keeping their students interested in completing their courses (Roper 2006).

4. Profit or Non-Profit? Funding and Revenue

Ideally, the game would be developed as a non-profit project funded by a research body, institution, or third party sponsor. If that were impossible, then the game could be developed on the basis of its future revenue. This could be achieved by the sale of the game either online or in high street shops. A game or visualization that ran on a mobile device would be an excellent way to monetize a project with an app costing a small amount, such as $0.99. This small starting price can still generate large overall revenue due to the enormous size of the mobile market. The other option is to fund the project through advertisements, which would show up in the game. These, however, would have to be carefully chosen to comply with the ethical boundaries of the project. There are advertising networks in place for iOS and Android that allow easy integration of advertisements into an app. The third option would be to sell players virtual money with which they could pay for services and buy goods in the game. The choice of funding for the development and licensing of the game would depend upon the needs and resources of the client and the project.

5. Conclusions

Show me your children’s games, and I will show you the next hundred years.

—Heather Chaplin and Aaron Ruby

It has long been understood that computer games can be useful tools in promoting teaching and learning. Playing a computer game can be rewarding, a benefit already recognized by educators two decades ago. Making games about archaeology using real ancient environments and modelled ancient personas will allow players to focus on details and landscapes, which can be looked at as long as they wish and explored at an individual pace. Exploration of a ‘real’ past world would also provide variety for gamers and such a game would create strong competition against the popular treasure hunting games, making players more aware of and attached to the ‘light side of the force’.

The benefits do not stop there. Archaeologists, historians, sociologists, economists, and other social scientists could find archaeological computer games a refreshing, novel and marvelous research opportunity. By studying and observing the ways in which the players of the game interact with each other and their environments, researchers can gain a greater insight into the ancient patterns of behaviour, everyday life, and decision making. Such studies can then also be compared with data from agent-based modelling and network analyses, especially fruitful in MMORPG versions which tend to give players the feeling that they are members of an exclusive community.

A game such as proposed in this paper can become an experimental ground for the testing of new and different social and economic theories related to the past societies. It also provides fertile ground in which the field of experimental archaeology may develop – in short, a huge brainstorming machine. The virtual environment, however, should not be confused for the realities of the past; although it can, as a partial representation, give a feeling of participation in the past.

The next step in the process is the creation and implementation of one or several concrete case studies. This would allow developers, players, and researchers to test systems and approaches in a continuing and interactive way. If such studies were framed correctly, they would also provide the basis for meaningful scholarly research projects as well as provide valuable tools for those working in heritage and education as well as for the general public.
Acknowledgements

The authors would like to thank the Oxford Centre for Maritime Archaeology at the University of Oxford, the School of Science and Technology at Nottingham Trent University, and the Alexander S. Onassis Public Benefit Foundation for their support. They would also like to thank Stefany Wragg of St Cross College, Oxford for editorial advice.

References


Re-reading the British Memorial: A Collaborative Documentation Project

Nicole Beale and Gareth Beale
University of Southampton, UK

Abstract:
The adequate documentation and dissemination of data relating to memorials in cemeteries and churchyards has been a constant challenge to researchers from community interest groups, academic institutions and bodies tasked with their preservation. The development of open source solutions for computational photography and the falling price of photographic equipment have placed many tools for the recording, analysis and interpretation of objects within reach of researchers from all of these backgrounds. By coupling these technological developments with a collaboratively authored and highly responsive methodology, the Re-reading the British Memorial project aims to improve and diversify the research practice of participants. The following paper will describe the unique blend of community archaeological practice, technical training and resource pooling which has characterised the project and will describe the way in which these elements contribute to its ongoing and sustainable development.

Keywords:
Community Archaeology, Polynomial Texture Mapping, Reflectance Transformation Imaging, Church Archaeology, Local History

1. The British Memorial

Memorials represent a unique but fragile record of the history of our communities. Almost every settlement in the United Kingdom has a collection of memorials which narrate the history of that community. Taken as a whole, memorials have the capacity to tell a much larger story. As well as allowing us to trace families and individuals these records also serve to illustrate a national history. They provide a unique and incomparably broad view of the social and economic changes which have characterised post-medieval British history and the lives of its people.

The unique value of memorials as a national archive has been previously noted (Mytum 2000; Willsher 2005; Commission for Architecture and the Built Environment 2007) and can be observed in the broad range of organisations and initiatives at national and local level which seek to document, preserve and catalogue these objects (Church Monuments Society 2002; Hintze 2011; Churches Conservation Trust 2012a). Of particular note have been the efforts made by voluntary community groups functioning at a local level. The efforts of these organisations have often been driven by the urgent need to record inscriptions which are endangered by weathering and decay and are often undertaken with limited resources. However, the popularity of the subject and the high numbers of volunteers who wish to assist in memorial recording means that progress is possible and is being made in many areas.

The Re-reading the British Memorial project aims to support these efforts through the provision of technical advice and training in cutting edge documentation methods and the initiation of collaboratively authored support and guidance documents for anybody wishing to record a collection of memorials. The goal of the project is to facilitate and support the development of a community of researchers for collaborative production of a methodology for memorial documentation and to assist in training others to utilise technical approaches in a sustainable way.

The aim of this paper is to explore the work which the Re-Reading the British Memorial...
Memorial project has undertaken up to now and to describe how this work may develop. The paper will place the project into context and will discuss some of the underlying principles and ideas which have informed its development. As the first published output of the project it is hoped that this paper will help to develop discourse around the idea of collaborative digital documentation in cultural heritage which will in turn refine and develop the project as it continues.

2. Existing Work

The project presented here grew out of and was inspired by a range of existing work in the fields of local and national history as well as in the development of digital and non-digital recording techniques.

It has been acknowledged for some time that changes in land use and also in communities and the effects of weathering and pollution have endangered many monuments within British cemeteries and churchyards (Commission for Architecture and the Built Environment 2007). The potential that these memorials may be lost without first having been documented prompted this project to look at current practice for the documentation of memorials and to consider how techniques utilised within the work and research being carried out by project partners, and legacy data held by the organisations charged with caring for the memorials might augment existing practice.

Several recent documents have been produced with the intention of guiding those wishing to record memorials of different types (Mytum 2000; Willsher 2005). These guides offer highly specialised advice on the analysis and interpretation of cemeteries and graveyards and represent an essential companion to any recording project. Also of interest to the project have been recent attempts by national organisations, including funding bodies and historical organisations, to encourage community involvement in the study and restoration of local cemeteries and historic religious buildings and to engage with existing community organisations (Heritage Lottery Fund 2012; NADFAS 2012; National Churches Trust 2012; Churches Conservation Trust 2012b). As a result of these initiatives the Re-reading the British Memorial Project has been able to actively collaborate with both local community groups and also with larger national organisations including Wessex Archaeology and The Churches Conservation Trust.

Techniques which were identified by the project to be of particular likely use within the context of memorial documentation include digital and computational photography techniques, such as Reflectance Transformation Imaging (RTI) as well as handheld GPS survey. RTI, developed by Tom Malzbender at Hewlett Packard Labs (Malzbender, Gelb et al. 2001), is a computational photography technique which uses a series of photographs taken with fixed camera position, exposure settings and focal properties. Within each image the light source is moved in order to simulate a dome-shaped lighting array perpendicular to the subject. These images are then compiled into one interactive file within which the object can be interactively relit, and settings such as light position, number of lights and rendering algorithm can be altered.

RTI has been at the heart of the project, and as one of the lower cost solutions introduced to groups by the team, it has proven invaluable both has a method of documenting and as a way to interpret memorials. RTI has been widely employed in the cultural heritage sector in the past few years, and has been used by a range of projects in order to analyse, interpret and disseminate objects and documents of historical significance (Mudge et al. 2006; Rabinowitz et al. 2009; Earl et al. 2011). Of particular note amongst these projects has been the recent RTI for the Study of Ancient Documentary Artefacts (RTISAD) project which has sought
to investigate the potential of this technique as a means of analysing a wide range of materials from the ancient world and has demonstrated the great potential utility of this technique for the interpretation of inscriptions (Earl et al. 2011). Figure 1 shows the method Highlight RTI in action.

The advent and subsequent boom in computational photography has been driven by the availability of open source and freely available technological solutions and also by a reduction in requirements for expensive recording equipment. Organisations such Cultural Heritage Imaging (Rabinowitz et al. 2009) and projects such as RTIiCAN based at Queens University in Canada (Gabov and Bevan 2012) have proven that RTI can be easily learned and employed by community groups with relatively little training and very few resources. Similarly, the advent of small scale handheld GPS has put relatively accurate GPS technology into the hands of most individuals and organisations, resulting in the surveying of cemeteries becoming less time consuming and more achievable than it previously has been.

3. Openness

At the core of the Re-reading the British Memorial project is an open methodology. Alongside the use of predominately open source solutions, this project is also attempting to use from the outset open approaches to project design and methodology development.

This project aims to use open design principles for objectives, workflow and methods development and testing. The development of the methodology is using an approach similar to that of the Detection of Archaeological Residues using remote sensing Techniques (DART) project (Beck, 2012) which is committed to open science principles, as this project intends to make the methodology available for use under open license and to work to eventually archive data with additional repositories.

Although this project uses many techniques to record sites, the technology that is at the heart of each survey is that of highlight based polynomial texture mapping (PTM) which is an implementation of the process of RTI. The RTI methodology being used by this project was developed at Cultural Heritage Imaging, and the method used by this project relies on open source software RTIBuilder developed by Barbosa at Universidade do Minho, and RTIViewer, developed by the Italian National Research Council’s (CNR) Institute for Information Science and Technology’s (ISTI) Visual Computing Laboratory. Both softwares are available under GNU General Public License. The major software for this project is therefore open source, and this reflects the overarching methodology.

In addition to the use of open source solutions wherever possible, the project also takes a community archaeology approach to decision-making and resource development. Community archaeology engages communities with archaeology from the beginning of any project, and ensures that there is an on-going dialogue between the community and the archaeologist (Simpson 2008, 5). This project uses the framework identified by Faulkner as ‘archaeology from below’, by avoiding the standardisation of archaeological investigations for standardisation’s sake (Faulkner 2000, 28). In Figure 2 members of the Branscombe History Project discuss RTI results after a day’s
surveying. The overall project has multiple smaller projects within it which are community-instigated and driven, with the major project team supporting activities, not driving them.

For instance, the project recognises from the outset that all stakeholders have multiple and varying requirements of the project itself, and the nature of the level of contribution of each partner is therefore not only negotiated at the start of each relationship, but is continually revised based on evolving needs. So far, the project recognises the following requirements, and it is anticipated that more will be added as the work continues:

- Lifelong learning
- Artefact conservation
- Digital repatriation
- Fundraising for sites
- Personal historical research
- Collective / community archaeological / historical research
- Advancement of knowledge
- Dissemination of knowledge

4. Project Update

In the instance of this project, community engagement manifests itself in numerous ways: Firstly, the project team is made up of a partnership between academics and professionals drawn from many disciplines, including archaeologists and historians, and special interest group members with a range of specialisms and expertise. An extensive period of partnership development and technology showcasing has been carried out since October 2011 to ensure that all parties involved have come into contact with both the technologies and the physical monuments.

Organisations involved in the early stages of the project have included The Branscombe Project, The Churches Conservation Trust, Highfield History, Portsmouth Royal Garrison Church, The University of Southampton, Exeter University and Wessex Archaeology. Each of these groups has been actively involved in refining the recommended recording practice and has participated in the transferal of skills and resources facilitated by the project. Until now project partners have been invited to participate whilst in latter stages of the project an open call for participation will be issued.

Workshops have been held at various sites in the UK to carry out recording of monuments, whilst simultaneously training team members to use the equipment and to recognise key characteristics of the artefacts being recorded. This skills swap has resulted in all partners having experienced both PTM and working with church artefacts.

The tangible outputs of these initial workshops have been the production of nearly one hundred RTIs and two site surveys. The guidance document for others wishing to engage in similar recording, which is based on consultations held during workshops, is currently being prepared for release as a wiki on our website, at which point it will be available for free editing by project partners. This takes its inspiration from projects trialling collaborative de-centralised strategy design approaches, such as the Smithsonian Institution’s open wiki for the design of the organisation’s new media.
strategy (Edson 2010). In addition to these outputs the project has also been able to initiate contact between partner institutions, a process which will be actively encouraged through further workshops and online platforms for discussion. This will help to establish a decentralised structure for communication, support and collaboration from the outset. As the project reaches its second stage an open call for participation will be issued.

Secondly, the project maintains a public blog (ourti.org) (Fig. 3), where all activities are documented. This site will also host all results as the project continues, and eventually will also host the methodology wiki.

Thirdly, the methodology design takes the form of a dialogue which began with initial training exercises at various sites, and will be formalised during a workshop scheduled in the summer of 2012 to which all project partners and stakeholders are invited.

Finally, expected outputs of the methodology design are the production of three-part guidance documentation consisting of a) a reusable methodology for the use of special interest groups and archaeologists who wish to collaboratively work to record graveyards using freely available and low cost technologies, b) technical guidance notes for the use of a variety of technology solutions, c) data management and dissemination guidelines. All documentation produced will be made available under appropriate open licences for reuse and redistribution.

The project invests much energy into providing real world impact activities for all stakeholders. Particular effort is spent on providing information for project participants who are not academic archaeologists. Social media is used to broadcast project activities, but also to gain support from other parties and to identify additional partners, and the team attend various public engagement opportunities to showcase the possibilities of PTM for recording cultural heritage.

5. On-going Work

The future usefulness of the data being produced by this project must be safeguarded, and therefore the importance of collecting and releasing data in structured datasets (Kintigh, 2012) is recognised by the project, and plans are underway to negotiate archiving of PTM data with the Archaeological Data Service (ADS).

The collaborative aspects of this project outlined above are representative of the theoretical approach, and alongside the open nature of the technical aspects of the work, this combines to illustrate how openness can be a focus for coordinated action. The methodology design is an iterative process, and the collaborative nature of this first phase of the project is an essential part of this.

Currently, the project is preparing plans for the next phase; the implementation of a web-based interface for the PTM database of memorials. From this it is hoped that data will be enriched by crowd-sourced contributions to documentation of major tell-tale information for identification of particular details, such as games to recognise particular symbols relating...
to religion, or to highlight details that point to conservation requirements such as lichen, cracks, tree roots, and stone tilting, and lastly for transcription from the images. GalaxyZoo’s work with the National Maritime Museum on the Old Weather project (Oomen and Aroyo 2011; Lang and Rio-Ross 2011) could represent a model for this next stage, and the team are looking at projects such as RunCoCo (Showers 2010) and Transcribe Bentham (Terras 2010; Causer et al. 2012) to identify possible methodologies for the successful integration of crowdsourcing into the project.

6. Conclusions

The Re-reading the British Memorial project has been well received by the communities within which it has been working up to this point. The initial consultation process has now been completed, and the project has been awarded funding by the Digital Humanities at the University of Southampton to host a day-long workshop to include a round-table discussion which will result in a formalisation of the methodology draft for release as a public wiki. The project process is iterative and therefore there is a need for informal and responsive support structures provided through on-going communication and networking of different organisations and individuals. It is anticipated that more partners will join the project as the work continues, and the team aims to remain open and inclusive throughout. Next steps include the formalisation of a multi-tiered partnership agreement, so that organisations and individuals are recognised more formally, and to ensure that resulting data can be shared using the appropriate licenses.

Community archaeology methodologies were the inspiration for the way in which this project has grown sustainably through a process of dynamic relationship building. Multiple organisations, groups and individuals have been involved in decision making and work package directions from the outset, resulting in all partners working together as mutually assisting. The project will now continue to grow and will modify organisational structure as it expands and develops with openness at its heart, through the collaborative creation of reusable resources and a sustainable training programme for all who wish to participate in the recording, interpretation and dissemination of our churchyards and cemeteries.

References


