IST-2002- 507382

EPOCH

Excellence in Processing Open Cultural Heritage

Network of Excellence

Information Society Technologies

D3.7: Final EPOCH Tools Description

Due date of deliverable: 15 March 2008
Actual submission date: 15 April 2008

Start date of project: 15 March 2004
Duration: 4 Years

University of Leuven

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

An infrastructure for the EPOCH tools

The goal of the jointly executed research in WP3 is to provide the infrastructure and the tools to build cultural heritage applications from. The paper on page 4 situates this research in the cultural heritage domain and shows the lessons learned during this project.

As illustrated in the information processing view of the process (cf. Figure 1), there are two distinct parts in solving this problem:

- First of all you need an appropriate set of tools to operate on the cultural heritage data. These tools produce data, ranging from text input over sound to video and 3D scanning; they consume data, e.g. for display, or they modify existing data, by enhancing and/or synthesizing the data. Chapter 2 describes the tools which were developed by Epoch partners, grouped by topic. Potential users find a short overview of each tool, giving its features and availability, in chapter 3.
- The way to exchange the data between the tools is through the information store. The paper on page 9 proposes an implementation of such a cultural heritage information store based on an existing content management system.

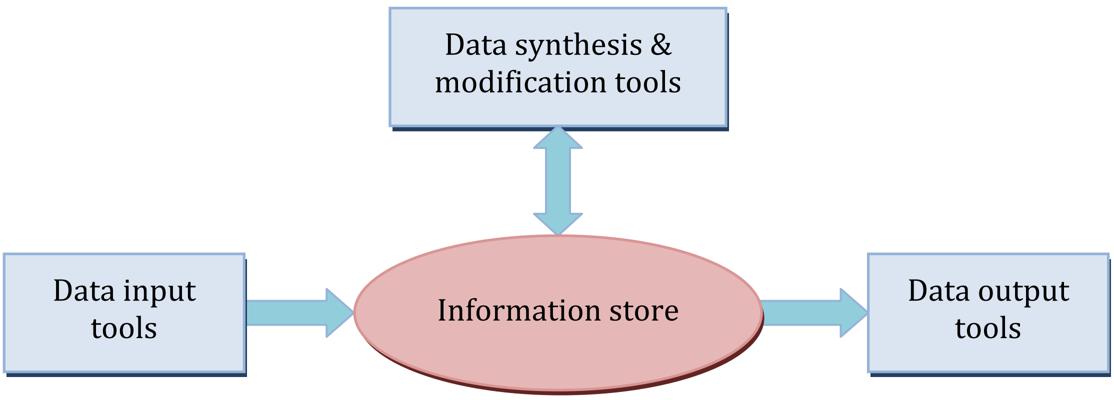


Figure 1: Information processing view of the cultural heritage processing pipeline.

EPOCH Conference on Open Digital Cultural Heritage Systems (2008)
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

The Development of WP3 in EPOCH

Sven Havemann[1], Luc Van Gool[2]

[1]Institut für ComputerGraphik & WissensVisualisierung, TU Graz, Austria
[2]Center for Processing Speech and Images (ESAT-PSI/VISICS), K.U.Leuven, Belgium

Abstract

The work package 3 of EPOCH, Jointly Executed Research*, is concerned with the analysis, research, and development of digital tools for the field of Cultural Heritage (CH). This application domain is very demanding for a number of reasons that are outlined below. The use of information and communication technology (ICT) is still immature. While in many other industrial and research domains completely digital workflows are common today, this not the case in CH. A number of proprietary or research solutions exist, but there is no coherent integration, either of standards and procedures, nor of software and data structures – which is fatal for the central goal, sustainability. During the EPOCH project, a number of factors contributing to this phenomenon have become much clearer, which was itself a major technical result of the project, since their identification permits the definition of strategies to overcome*

1. Introduction

Perhaps maybe the greatest achievement of the EPOCH project on a technical level was that it brought together specialists from diverse fields, made them collaborate on technical problems, and thereby fostered a much deeper interdisciplinary understanding of the diverse research problems in the different fields. A concrete development roadmap for a reasonable application of ICT in CH could not have been devised when EPOCH started; but it is much clearer now. This is the real achievement, since it represents a genuinely interdisciplinary effort in an environment where many cultural heritage professionals were justifiably suspicious of the promises of technologists whose track record did not encourage confidence that the exotic tools proposed would either mesh smoothly with the sectors workflows or be sustainable in the longer term. Technological developments have in fact been littered with examples of “neat ideas”, developed as research exemplars but never achieving robust deployable solutions and rapidly becoming obsolete. Cultural Heritage is a very demanding domain for the application of ICT for a whole number of reasons:

First of all, Cultural Heritage is a very demanding domain for the application of ICT for a whole number of reasons.

- **CH is a broad field**
  CH ranges from the archaeologists digging in the excavation field to the historian developing interpretations and building hypotheses about historic facts and relations, to the museum curator who chooses from the available information to set up an exhibition. Other main actors come from engineering, also from diverse fields: From photogrammetry (geodesy) over computer vision to 3D scanning professionals, but also from semantic technologies, to user interface design, to interactive AR/VR, and many more. But besides these main fields more specialized applications exist-as examples consider the assembly of fragmented objects, and the particular requirements of underwater archaeology.
- **Archeology is a desctructive science**
  In order to get access to strata with older remains deeper in the ground, the strata from younger periods have to be removed. The original information from the younger strata will never again be accessible: each period of Troy can only be excavated once. This puts great emphasis on faithful recording and documentation. The problem is that it may become apparent only much later that a particular piece of information would have been of great value. Therefore, in principle, as much data as possible from each stratum must be recorded and sustainably stored âĂŞ even data that may seem irrelevant at the time. CH professionals are understandably reticent about experimenting in this one-off opportunity with technologies whose effectiveness is not thoroughly tried and tested.
- **Overwhelming masses of data**
  Excavation sites comprising old villages and cities are typically quite large in extent, typically ranging from several hectares to a few square kilometers. Faithful recording, e.g., laser scanning at millimeter precision over the whole site, is not feasible, not only because of the data acquisition cost, but also because it requires extremely efficient data mining techniques to avoid the data grave problem: *Data that can not be accessed are useless data*. Some parts, e.g., bas reliefs carved on a block of stone, may require sub-millimeter precision. Extreme examples, like the huge site in Kharnak (Egypt) or Angkor Wat (Cambodia), sites may be literally covered with stone fragments carrying carvings and extending to several square kilometers. Balancing accessibility, accuracy and storage requirements remains a problem.
- **CH institutions are not rich**
  Digital technology is coming down in price, but acquisition de-

vices such as laser scanners are still quite expensive, and even digital cameras are sometimes considered too brittle for the field. The highest cost, however, is that associated with providing trained, specialized personnel for the digital equipment. In addition quite often, the post-processing cost can be higher than the acquisition cost. New devices such as camera domes have a huge potential for the automatic acquisition of smaller findings, but still, much research is needed to develop more automatic approaches suitable for use by staff who do not have or require the same level of technical competence and training.

- **CH professionals are conservative – for good reasons**
Specialists in ICT sometimes find themselves confronted with seemingly irrational arguments from CH professionals, who deny the usefulness of digital technology altogether. The main archological recording tool is pencil and paper, training the drawing skills is an important part of the education. But sketching allows interpretation, while taking a photo does not. Photo-realistic renderings of the past, like in movies a la *Gladiator*, all too often make it impossible to distinguish between fact and fiction. Computer-based tools for interpretation of digital data are typically more difficult to use than pencil and paper, but are currently still not more objective. So what is their added value?
- **Sustainability is a prime concern**
It is without a doubt a problem when a historic artefact, that was lying safely in the ground for 2000 years, is digitally recorded, but the digital record becomes unreadable within 20 years. Either the data carrier becomes unreadable (*physical degradation*) or no software exists anymore that can interpret the data (*format degradation*). To address the latter, well-documented data format standards need to be developed. This, however, is more difficult than it seems, because practitioners will only cease to use ad-hoc or proprietary formats if the standards, although they are fixed, are flexible enough to cover all real use cases. And extending standards is not a solution either, since blown-up standards can become a problem themselves.
- **Documentation and interpretation**
There has to be a clearer distinction between the two tasks of CH: Faithful recording and documentation, is an engineering task (taking measurements), whereas interpretation of the data requires a huge background in culture and art history. Formally, the act of interpretation is to establish, by scientific reasoning, semantic relations between pieces of data. ICT could help greatly to make the interpretation process more transparent, as it allows to establish the relations explicitly, rather than only implicitly.
- **CH – human science or engineering domain?**
EPOCH technical work has made clear that the deficit is caused by both sides: CH professionals should be more technologically aware, and should actively try to find the best solutions available. Equally important, however, is to spot inadequate solutions and missing technology. Technologists are quite often too naive and offer for complex problems technical solutions that are either too simplistic, or too cumbersome (and sometimes even both).

2. A Plethora of Applicable ICT

On the one hand, the use of ICT in CH is still premature. On the other, however, a considerable amount of research has been invested to develop technology that can help. The range of technologies investigated in EPOCH gives an impressive view on all the different ingredients, many of which are covered in-depth by articles later in this report. This not only illustrates the great potential for the application of ICT in CH, but also gives an idea of the complexity of the task to develop integrated solutions.

- **Semantic technologies (CIDOC-CRM)**
Raw data (measurements) are useless without meaning. CIDOC-CRM is a technology for creating semantic networks [CDG*05, Doe05]. It defines a set of *entities* (84 classes such as place, time span, man made object, information carrier) between which *properties* can be established (141 relations such as participated in, performed, at some time within, took place at, is referred to by, has created). Obviously, CIDOC-CRM is a *horizontal technology* meaning that any piece of software in CH should ideally implement it to maintain the meaning of the data when processing them.
- **3D acquisition technology**
Laser scanning is all too often considered to be a solved problem and to work "out of the box" – which it does definitely **not**. All groups and individuals trying to apply scanning technology go through the painful experience that a less than optimal planning of a scanning campaign, with lapses from unsuitable file names to not recorded acquisition positions, can increase the post-processing time by arbitrary factors. Several contributions of EPOCH members, e.g. at the VAST conferences, have shown that this problem is taken seriously, and that robust solutions are under development.
- **Measuring methodology (ref to geodesy)**
Measuring is difficult: The truth is that in general a measurement is never exact since it is only one sample from a whole probability distribution. The probability of taking a sample that lies close to the true value can be adversely affected by many factors. To guarantee a certain precision for the measurement (again only with high probability) requires two things: First, a mathematical model of the measurement process (including technical devices and environmental conditions), and second, to calibrate all processing stages. Calibration before measuring is unfortunately not common practice, nor is a culture of *measuring always twice*, to acquire more than a single sample from the probability distribution.
- **3D shape processing**
The easy manipulation of 3D objects, amenable also for non-expert users, is unfortunately still an open problem. Under the hood of EPOCH, the Pisa group (Cignoni/Scopigno) has produced the Meshlab tool as one answer to this issue. The idea behind it is the well-known Photoshop metaphor, so that shape manipulation resembles very much to image manipulation tools, which more users may be familiar with.
- **Seven Open 3D Research Problems**
Some more fundamental issues with current 3D technology as a whole have been identified by Havemann and Fellner in [HF07]. These are (i) to classify all possible 3D shape representations, (ii) the 3D file format problem, (iii) stable markup of 3D parts, (iv) representation-independent 3D query operations, (v) to document reliably the provenance and processing history of shapes, (vi) to close the semantic gap, and (vii) to maintain the relation between shape and its meaning consistent.
- **Long-term preservation standards + metadata**
There are lots of competing (national) documentation standards for museums, but only few in fact for archeology; in particular, none that is commonly accepted. The technical device to translate from one documentation standard to another is an ontology. Ontologies are an active field of fundamental research in

computer science, often quoted in the context of the semantic web. Upcoming standards from the library scene, such as FRBR [Til04, K.G98], show a shift from old-fashioned Dublin Core to semantic technologies. All these are being harmonized using ontologies, only to mention FRBR-CRM.

- **Photogrammetry and Computer Vision**
Old objects of cultural value typically have complex materials, which makes them difficult to scan. On the other hand, history has left its traces on their surface, and high-frequency variations are just perfect for matching stereo pairs. The potential 3D acquisition from simple photographs, taken in a systematic manner (*image sequences*) has been demonstrated by the EPOCH reconstruction webservice in an impressive way.
- **Camera domes and other novel acquisition devices**
The more controlled the image acquisition is, the more precise is the 3D reconstruction from photographs. Taking pictures with systematic variation of the incident light direction allows faithful rendering of the appearance under arbitrary lighting conditions (*relighting*), and to approximate the BRDF in each surface point using *polynomial texture maps* (PTM) or *bi-directional texture functions* (BTF). This principle can be exploited using various technical setups: single/multi-camera, single/multi-light, with/without mobile parts, immobile/transportable. This technology also has the potential to be used for the *mass digitization* of artefacts.
- **Procedural modelling tools to build hypotheses**
Destroyed or decayed historical sites can not be scanned anymore, and the physical remains are typically incomplete and deteriorated. Most often, a complex puzzle needs to be solved in order to come up with an appropriate hypothesis about the past. But how can CH professionals iteratively develop, exchange, assess, and refer to highly detailed three-dimensional representations of the past state? Procedural 3D modeling has the potential to greatly facilitate the construction of complex objects by certain rules. Two such technologies were explored in EPOCH, namely shape grammars (CityEngine) and generative modelling (GML).
- **Content management systems for CH**
Current database technology and content management systems are good at managing textual data. Content-based search in text documents (pdf) works quite well, but CH heavily relies on digitized artefacts, images, sketches, or various sorts of 3D. Different technologies for managing multiple types of data have been investigated in EPOCH, but no ideal candidate was identified. One problem is that 2D and 3D data are mainly treates as anonymous BLOBs (binary large objects) by current DBs and CMSs.
- **Avoid the data grave problem by advanced search tools**
Data that can not be found are useless data; it is as if they never existed. Efficient search capabilities are of prime importance. A pre-requisite of searching, however, is indexing: Good solutions exist for extracting the characteristic words in a text document. Image retrieval works reasonably well in some domains, but how to best search 3D objects is still an open question. EPOCH has deliberately chosen to leave this area of research to more specialised projects. In the future, however, ICT applications in CH will not succeed without decent 2D and 3D searching capabilities.
- **Find solutions for IPR issues**
In principle, European history belongs to the people in Europe. But to dig out historic artefacts is tedious and costly, as well as to tidy, collect, register, and store them. Archeologists and museum staff must be paid, as well as buildings and exhibitions. Who holds the rights of a digital artefact, the owner of the artefact or the scanning company? How are revenues are split from derived works, such as CD-ROM sales, internet download, or even physical replicae? And in either case, how can the IPR holder exercise his rights? A vast number of legal issues still requires solutions, which should be taken in a way that a novel market of digital historic artifacts can emerge, and create revenues for all involved stake holders.
- **Analysis and interpretation tools for legacy data**
Masses of pre-digital historical information exist. Old history books, since closer to ancient times, sometimes contain extremely valuable information. Not just the digitisation is an issue. More serious is that although old catalogues are valuable sources of information, their systematics and their terminology are outdated and incompatible. The modern view on complex information is that of a semantic network with entities and relations. Fact and interpretation need to be linked in a bi-directional way. Consequently, tools need to be developed for semantic data mining, to exploit and preserve the old sources in an optimal way.
- **Consequences of the London Charter**
Being able to assign a meaning to pieces of data is a complex issue, since data are frequently transformed and processed. The London Charter of Beacham, Denard et al. [Lon06] addresses this problem by defining general rules and guidelines. It introduces the important novel concept of *paradata*, which comprise all information that tell about the provenance of a dataset. This is also a horizontal requirement, since all tools in a processing tool-chain have to support and add to the processing log.
- **Guided tours through museum exhibitions**
Historical museums must compete today with many other interesting leisure activities, from science museums over edutainment exhibitions to featured tourist attactions. Also historical movies and computer games have contributed to raising public expectations to extremely high levels. The classical museum, with long textual explanations for each exhibit, is hopelessly outdated for the new generation of media-aware, unpatient visitors. Audio guides are standard, but are only a step on the way to personalised museum exhibitions. Techniques such as RFID tags and interactive displays (*multimedia kiosk*) are on the rise – all waiting for high-quality, authentic digital content that must be conveyed in an engaging, easily understandable fashion. First steps in this direction were taken in EPOCH. But more and better tools are needed to let museum staff create, in an efficient way, such novel types of combined exhibitions integrating both digital and real content.
- **Multi-modal user interfaces**
The success of the Nintendo Wii has shown that unconventional user interfaces using, e.g., gesture recognition, are suitable for a mass market. The public audience understands and demands such devices. More and more camera-based tracking techniques are available, users will soon expect ambient technology that is able to react. A number of interface techniques were investigated in EPOCH. Psychological research shows that knowledge that is actively explored leaves much deeper traces in memory than the passive consumption of an information stream. The internet stimulates average users to browse and explore information spaces, the same will be expected from museums – only better.
- **Avatar technology**
Maybe the best way to engage people with history is through tales and stories. Telling about people's everyday life in the past can convey a very direct and intuitive understanding of human condition in acient times, and can motivate and contextualize the resulting developments. But stories require actors for the audience

to identify with. EPOCH has actively stimulated the proliferation of avatar technology, first to populate otherwise void historical 3D reconstructions, and second, laying the foundations to avatar guided tours even providing limited conversation capabilities.

- **Dissemination to end users: Engage web communities** Community-based technologies (*Web 2.0*) have great potential to unleash a great momentum, especially with authentic historic content: The idea is to engage interested end-users by a mixture of a history-focused user-generated semantic network (a la Wikipedia), and rich 3D interaction with historical surroundings, a la SecondLife, but with a *time slider*. With increasing momentum, CH-professionals can no longer ignore it, and will gradually take over to assure the historic and scientific quality of the content. This idea came up during EPOCH, but was not explored within the project.

3. The Pipeline View versus the Repository-Centered View

Originally the idea of the technical work in EPOCH WP3 was to start from a technology review to see which sorts of software tools are available, which are missing, and then to complete the typical CH workflows. After the initial requirement analysis the project has set up an internal call for **New To**ols **N**eeded, or short *Newtons*, to close the identified gaps.

In principle this worked quite well in that the Newton projects did good research and produced an interesting suite of tools, as explained in detail in a number of VAST proceedings. However, one implication of the fact that use of ICT in CH is still premature is the lack of established digital workflows. One of the few notable exceptions is the production of *multi-lingual audio guides*, routinely used today in many museums and cultural attractions as a comfortable way of providing museum visitors with a flow of information. The audio guide production workflow is now completely digital (information from Daniel Pletinckx).

Another idea that was very influential in EPOCH from the very beginning was the **CH processing pipeline**, roughly in three stages:

- **Excavation sites** are digitally documented through scanning and photographs, interesting artefacts are scanned as well and are put into a database, and semantic and meta-information (when, where, who...) is faithfully added
- **Scholarly research** of historians and archeologists seeking for interpretation of the found evidence adds further information; fact and interpretation are linked bi-directionally, the semantic network emerges
- **Museum exhibitions** finally spread the acquired historic knowledge and disseminate the past to the public; museums are open for generating additional revenues by novel digital distribution channels; but museums are also the guardians concerned with the long-time archival of many of the most precious items

This pipeline view was initially taken too literally. We thought the job of the people in the field was only to record and thereby produce digital data in the first place (in reality they are less interested in recording than in interpreting the facts); that this data was, through a sequence of processing steps, refined and made consistent, until it is finally displayed. Naive was to assume that we would only have to fill in the gaps in the sequence of processing steps in order to improve the situation. We thought it would be sufficient to define a number of digital exchange formats, and create a sequence of software tools, one importing the results of the previous processing step, do some processing, and then exporting to feed the next step.

It is not as simple as that. The truth is that we finally found that the most annoying lack of technology was that we had no central information storage in EPOCH. This area of research was left out because the development of a CH repository infrastructure was exactly the objective of the BRICKS project. However, it turned out that there is a subtle mismatch: BRICKS is for reliable archiving of masses of data. What EPOCH needed was a central information store that could keep track of the different pieces of data in the various stages of transformation. This would be the place to reliably store, for instance, the information that scene A contains objects B, C and D, where C was produced from scans C_1, C_2 and part of C_3, but C_3 was actually the result of applying the hole-filling algorithm X with parameter set $(p_1, p_2, \ldots, p_{27})$ to the raw scan dataset C_3'.

The refined pipeline view at the end of the project is that there are data producers, filling the digital library with substantial amounts of raw data, there is data refinement, taking from and feeding into the library, and there is browsing and searching to take data out of the library, e.g., for public presentation. In any case, the digital library is the central facility.

4. Lessons Learned: CIDOC-CRM and the London Charter

One of the greatest assets of Cultural Heritage is its ***authenticity***: Our past is something that matters, that each and every individual can directly relate to. There may be different and even conflicting views on history, but there are also undeniable facts and findings. Most crucial for taking benefit from this asset is that authenticity is not jeopardized by mixing up fact and fiction, truth and assumption, invented and real.

Two so-called *horizontal technologies* were mentioned in this text, CIDOC-CRM and the London Charter. The trouble with horizontal requirements is that they can be very fragile, since if one tool in a processing chain does not respect them, information may be destroyed or become inconsistent. It is not easy to define and implement paradata, for instance, in a way that is robust against accidental damage: Assume that for a triangle mesh dataset an annotation file exists that faithfully documents, for each single triangle, whether it was really measured, or whether it was produced by some hole-filling algorithm in a post-processing step. Note that this information can be crucial: There are very critical subjects in the CH community who have declared that they deny drawing any scientific conclusions from a 3D dataset with unclear provenance. Is there really a cuneiform writing on the wall, or is it only invented by some algorithm? In other words, if paradata are lost the dataset becomes useless and can in fact be deleted. – Well, the first step to avoid such scenarios is to introduce a strict *write once*-policy: Never save a dataset, when processed, under the same file name as it was loaded.

This example shows that both horizontal requirements we have identified in EPOCH to be of great importance have large implications for standards and procedures, and in particular for software design: the London Charter with its notion of paradata and of *intellectual transparency* (to always be able to discern fact from interpretation), as well as CIDOC-CRM with its idea that each and every piece of information, fact or fiction, object, place, or time span, is embedded in one large common cultural information space – and that these relations should indeed also be explicitly represented in the data.

One great result of EPOCH is that researchers from various different fields got acquainted with both of these horizontal requirements.

They have started to consider and think through the resulting implications for their respective fields of research. The lesson learned is that this process should have started much earlier in the project.

5. A Plan for the Future

As mentioned in the introduction, it would have been impossible at the beginning of the project for the EPOCH team to devise a detailed plan for a large-scale technical project that substantially improves the situation and helps CH institutions to unleash the huge potential of digital artefacts.

Now, at the end of the project, it is clear that we first have to take care about the beginning of the CH pipeline: Crucial is to have masses of high-quality artefacts, acquired routinely with cost-efficient equipment and only moderate post-processing expense. Only with impressive content it will be possible to make 3D technology a standard, rather than only a gadget. To focus on, design or develop only dissemination activities, i.e., to create the demand at the end of the pipeline, is pointless if no high-quality content is available: content is the key. The commercial potential of stunning high-quality content will immediately become apparent, and without a doubt a new market for digital cultural artefacts will emerge.

A project that could bootstrap this development would most probably look like this:

1. WP: **A repository infrastructure for CH**
   All shape data, image data, metadata, paradata, semantic knowledge, and shape semantics need to be collected in a common information storage, a central digital library (DL).
2. WP: **Automatic devices for mass acquisition**
   We need a new generation of robust, reliable, automatized camera-based 3D acquisition devices that work, e.g., for small artefacts in a conveyor belt fashion: Take the artefact out of the box onto the belt, let it be scanned, and put it back into the box; and in a similarly simple also for larger artefacts.
3. WP: **Artefact processing with TLC built in**
   Conventional mesh processing software does not collect paradata. Mesh repair, e.g., with the high-end software GeomagicStudio, lets not tell afterwards which triangle was measured, and which was invented. We need to seemingly re-invent existing functionality, simply because to integrate our horizontal requirements requires substantial software re-engineering.
4. WP: **Creating the collection item, the digital artefact**
   Entering data into the DL is not as trivial as it seems: We want to collect paradata right from the beginning. Furthermore, semantic information is qualitatively superior to old catalogues following the DublinCore paradigm. So the ingest of legacy data always will be only semi-automatic, we need efficient tools for that.
5. WP: **Searching and browsing in 3D collections**
   Collecting data is pointless if it is not possible, or even just difficult, to retrieve the data. We have to investigate all different sorts of indexing strategies to offer the user a wide range of searching facilities. Browsing the DL must be very easy, and so must be to find relevant pieces of information.
6. WP: **Synthetic artefacts: Hypothesis building**
   We have to support researchers in systematic hypothesis building. Our tools shall offer easy ways to document decisions why a particular artefact might have been at this or that position in the past, or used for this or that purpose. It must be easy for CH researchers to create, exchange, and reason about 3D scenes.
7. WP: **Presenting history to the end user**
   Although this is an acquisition rather than a dissemination project, we may not neglect the requirements of display technology. In particular we have to create, for each shape and material representation acquired and supported in the project, a suitable high-quality interactive rendering algorithm.

The EPOCH team would just love to help making this vision become reality.

References

[CDG*05] CROFTS N., DOERR M., GILL T., STEAD S., STIFF M.: *Definition of the CIDOC Conceptual Reference Model*, version 4.2 ed. CIDOC Documentation Standards Working Group, June 2005. also ISO/PRF 21127 (almost). cidoc.ics.forth.gr.

[Doe05] DOERR M.: The cidoc crm, a standard for the integration of cultural information. (online), Nuremberg, 2005. Tutorial slides, cidoc.ics.forth.gr/docs/crm_for_nurnmberg.ppt.

[HF07] HAVEMANN S., FELLNER D. W.: Seven research challenges of generalized 3d documents. *IEEE Computer Graphics and Applications 27*, 3 (2007), 70–76. (special issue on 3D documents).

[K.G98] K.G. SAUR: *Functional requirements for bibliographic records: Final report*. Munich, 1998.

[Lon06] LONDON CHARTER INITIATIVE (HUGH DENARD): The london charter, June 2006. www.londoncharter.org.

[Til04] TILLETT B.: *What is FRBR?* Library of Congress Cataloging Distribution Service, 2004.

EPOCH Conference on Open Digital Cultural Heritage Systems (2008)
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

A Common Infrastructure for Cultural Heritage Applications

L. Van Eycken[1], B. DeKnuydt[2], and L. Van Gool[1]

[1]Center for Processing Speech and Images (ESAT–PSI/VISICS) – IBBT, K.U.Leuven, Belgium
[2]IMEC vzw – IBBT, Belgium

Abstract

This paper explores the infrastructure needs for cultural heritage applications. Small dedicated applications as well as very large projects are considered. A unified approach for managing cultural heritage information is proposed to handle storing and exchanging data. An implementation demonstrates its use in two different cultural heritage applications.

1. Introduction

The jointly executed research of the Epoch project aimed at providing the infrastructure and the tools to build cultural heritage applications from. Starting from stakeholder needs and implementation experiments, we looked for a common infrastructure to support the various applications. An important aspect of cultural heritage applications is their diversification. This relates not only to the kind of data (descriptive text, semantic relations, location information, sound, image and 3D representation, ...) but also to the way the data is used (preservation, interpretation, visualisation, exhibition, web catalogues, ...).

Let's consider as an example a large historic site, like the Valley of the Kings in Egypt. The amount of objects that need a description is enormous: there are about 627 tombs, with hundreds of noteworthy items (sarcophagi, masks, wall-paintings ...) per tomb. Each item itself usually contains many 'views' or 'presentations'; think of 3D scans of a mask, photographic pictures, X-ray data, textual descriptions ... Many of these views themselves consist of several data objects: frontal pictures of Tutankhamen's mask, dating from 1799 up to 2007. Of course, we don't know in advance all possible future applications, but our data model should at least support the following applications:

1. *Show/navigate the data*
   We want to be able to present and access easily all parts of the data, from the most general overview to the smallest detail. In this example it means accessing data from airplane flyovers of the terrain up to the tiniest inscription, from old black and white pictures to X-ray data. And be only a 'click away' to navigate from e.g. this inscription to comments from scholars on it and translations in modern English.
2. *Present the data in an exhibition*
   We must be able to distill self-contained subsets of this information for expositions, textbooks, museums etc. For example, distill sufficient material for an on-line exhibition of pictures from the excavation of grave KV21 in 1817.
3. *Preserve the data*
   We want to maintain the information for at least 100 years. This not only means preserving the data in the literal sense of the word, but also keep it accessible under the fast changing technological conditions, and update it with any new information that might come available.

Apart from the previous example with large data collections, there exist also many small applications in the cultural heritage domain which require only a limited usage of a confined data set. Due to this diversification, it's not possible to define one single infrastructure technology, appropriate to support all cultural heritage applications. Therefore, we propose a general framework as an architecture, which can be tailored to any real application.

The requirements, as summarised in the next section, define the boundary conditions for a cultural heritage architecture. Section 3 explains why we prefer the information processing view over the pipeline view as the basis for a common infrastructure. Section 4 shows how data objects can be used to group all related information in an efficient way. These data objects are stored in an information management system, as described in section 5. The paper concludes by illustrating the concepts in two implementation examples.

2. Requirements

The important considerations of the research on stakeholder needs for the definition of a common infrastructure are:

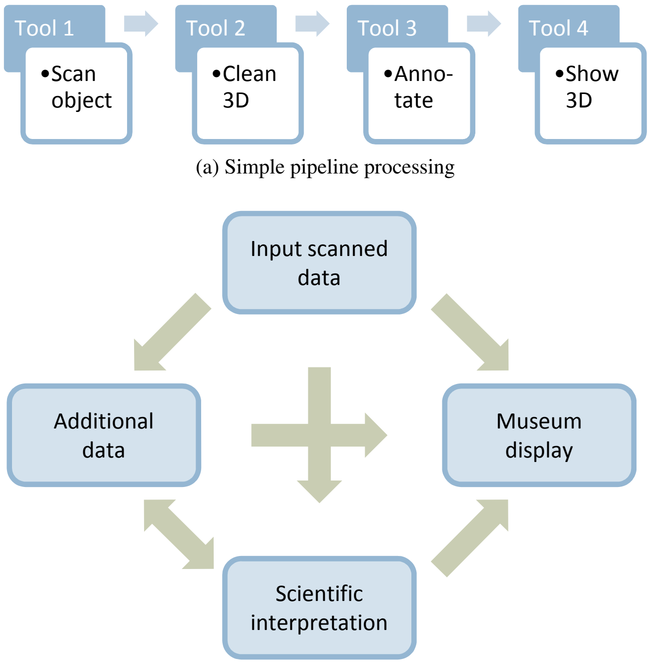
- need for a framework to guarantee interoperability among existing systems and facilitate data reuse;
- need for sustainability in regard to the accelerating progress of technology;
- need to maintain international standards;
- need to encourage an open source approach;
- attention to safeguarding scientific accuracy during data collection and processing by making a clear distinction between fact and hypothesis.

Good interpretation of cultural heritage also means that a wide range of sources (oral, written, research, traditions ...) is used to create the interpretation and presentation, as the result of a multidisciplinary study. This not only means that the framework is conceptually valid for all cultural heritage domains but also that it allows exploiting all available information sources. In technical terms, this means that its data structures allow information from multiple domains. Today, databases in cultural heritage are usually tailored to a single discipline and are not suited to contain multiple domain information, created by a multitude of researchers, physically present at different locations.

To ensure proper involvement of all parties concerned, interpretation technology needs to be anchored within the stakeholder communities, and not be the monopoly of high tech companies. In technical terms, we need an affordable and easy content management system as the kernel of interpretation systems, which must be designed to be adaptable and open.

From the technical point of view, two additional requirements are added. The framework must be as implementation technology independent as possible, since we want to prevent it from becoming outdated too quickly. This also implies that the framework cannot be restricted to the currently available tools and applications only. It should be able to embrace new emerging technologies.

To guarantee sustainability, the framework should be based on international standards as much as possible. The open source approach is also encouraged, but not enforced. Users should be able to use commercial tools and businesses should be able to develop based on a common infrastructure.



(a) Simple pipeline processing

(b) More complex application

Figure 1: Examples of processing in applications.

3. The concepts behind a common infrastructure

An application in general consists of a number of tools working on data. Each tool performs a well-defined subtask of the application. In practice the tools often form a sequence of operations: a tool works on the data generated by a previous tool and delivers its data to the next tool, as shown in figure 1a. This is commonly referred to as a "pipeline". On the other hand complex applications, such as the one in figure 1b can't be described as a pipeline: it consists of several sub-applications which combine and iterate upon exchanged data. This data can even be generated or used at different times and places.

3.1. Classification of tools and applications

There are several orthogonal ways to describe cultural heritage applications. A first dimension consists of the seven functional technology areas as described in the application for the Epoch network. These areas correspond to the numbered boxes in figure 2. This classifies cultural heritage activities from a computer science point of view.

A second dimension describes the chain of practice in archaeology and museums. This cultural heritage pipeline corresponds to the unnumbered boxes in figure 2. This figure also visualises which technology area impinges upon which part of the archaeological or cultural heritage process, but without the feedback loops and iterations. It is important to note that due to the loops and iterations the process is not a real pipeline where information enters on one side and propagates through transformations towards the end of the pipeline.

A third dimension enumerates the different tools for the different application domains. The domains correspond to (subsets of) the different stakeholder communities as described by the Stakeholder Needs team (Heritage Policy, Local Authorities, Cultural Heritage Sites, Museums and other Cultural Heritage Organisations, Associated Communities, Tourism, Education, and Technology). Tools may be specific for a domain or may be shared. For instance, avatars can be used in many applications but Harris matrix tools are very specific to excavations.

This classification is a good way to categorise tools, but it is less useful for defining a common infrastructure for cultural heritage applications. Such an application can be represented by a cloud of points, which cover only a part of this classification space. Even for applications with similar goals, their point clouds can be very dissimilar. On the one hand, this classification space is too extensive to derive a common infrastructure from it. On the other hand, each dimension in itself is not enough to describe all applications so it cannot be the sole basis of a common infrastructure. Therefore, we need another approach to tackle the problem.

3.2. The information processing view

Another way to look at the cultural heritage processing is to follow the information through the application. A first conclusion is that very few applications cover all aspects of data processing in a domain. Most applications deal with only one aspect or at most a few. A consequence of this fact is that applications are not self-contained: they use data from or produce data for other applications, so we need intermediate storage. This is certainly needed if loops and iterations are present, not inside but between applications. Furthermore, data is often re-used by different applications. A final conclusion is that many applications may take their inputs from multiple other applications.

For this kind of processing, the information flow can be depicted as given in figure 3. All the data is stored in the information store. This data can not only be used to generate output, but it can also be modified or synthesised into new data. Having a central point of information exchange has the advantage that it is easier to decouple

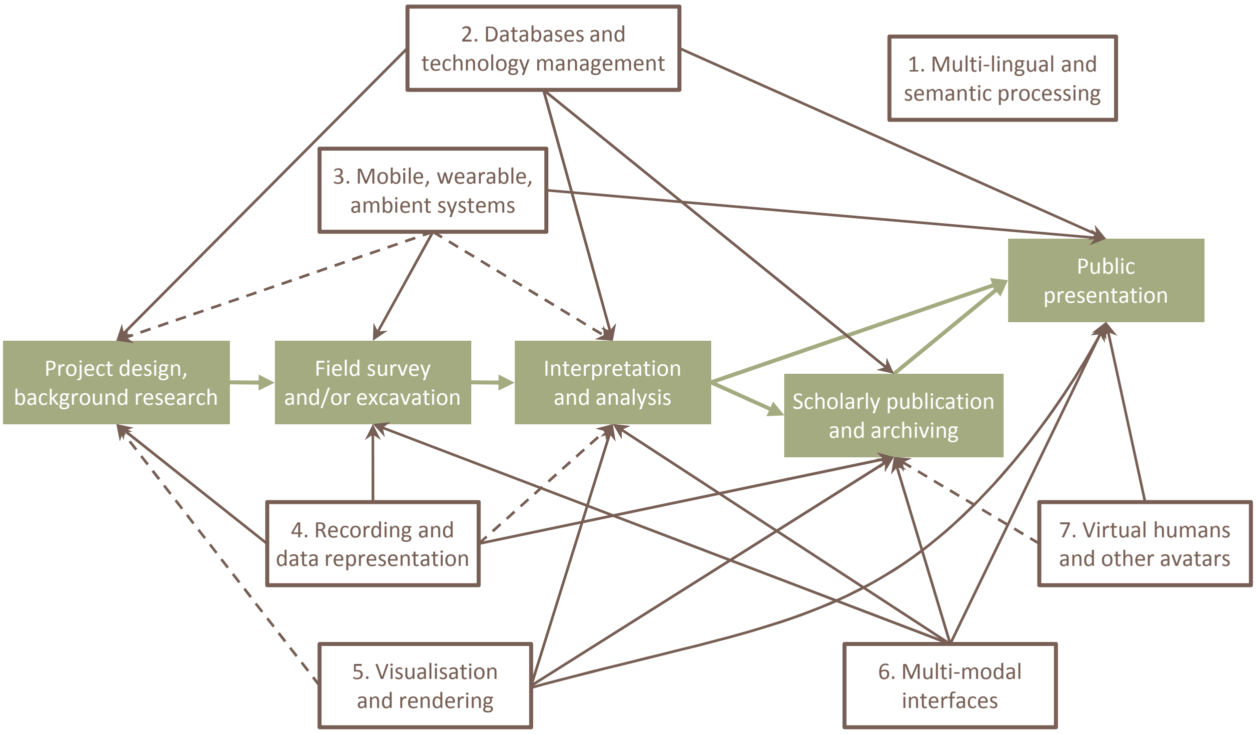


Figure 2: Relation between the typical cultural heritage pipeline (the filled boxes) and the functional areas (the numbered boxes) as proposed by Nick Ryan.

the generation, manipulation, and final use of information when needed, as requested by the stakeholders.

The central element is the information store. In theory, it stores all cultural heritage data, which was ever produced or which will ever be used by cultural heritage applications. In order to avoid mixing up measured data and interpretation, this store should hold the raw data (the original "mother data" as well as the transformed data) with the accompanying metadata, which describes all extra information and transformations applied to it.

Data is input in the data store in two ways. First of all, data producers generate data for it. Examples of such producers are tools to measure/scan objects and field data, or tools to add comments. Another way to generate data is to modify or synthesise existing data. The data synthesis implies that multiple sources are used to generate new data. Examples of these transformations are generation of virtual objects and worlds, annotations, and the interpretation of data. Finally, the data from the information store can be output to users

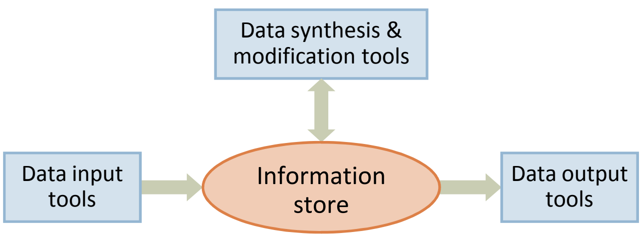


Figure 3: Information processing view of cultural heritage applications.

by data consumers. Usually they also use multiple inputs. Examples of data output are not only all kinds of display, with or without user interaction, but also the outcome of queries for information from the data store.

Of course, a real application cannot hold all cultural heritage data and all tools for working on it. In practice we only need a subset of the information store as well as of the tools, just enough to get the job done. If needed for other applications, we can merge it with other information stores. This is not a problem as long as all information stores are based on the same foundation and data synchronising tools are available.

3.3. Concepts of an architecture

As illustrated in the information processing view of the process (figure 3), there are three parts in solving this problem:

- First of all you need an appropriate set of tools to operate on the cultural heritage data. These tools produce data, ranging from text input over sound to video and 3D scanning; they consume data, e.g. for display, or they modify existing data, by enhancing and/or synthesising the data. Since many links exist between the different data items, we may need a way to keep all related items and their link information together as one data object. Section 4 proposes a way to achieve this.
- The way to exchange the data between the tools is through the information store. This information store does not necessarily correspond to a physical storage such as a database, but it may. For instance, when data is processed within an application in a single pipeline, the intermediate data can be kept in memory. But when different applications are using and exchanging information,

a more formally defined storage implementation may be required. Section 5 proposes an implementation of such a cultural heritage information store based on existing content management systems.
- Finally, exchange formats (shown by the arrows in figure 3) must be defined to guarantee a proper exchange of information, raw data as well as metadata. For long time preservation additional requirements should be met. Epoch is not proposing to define new or adapted standards. In our opinion, enough standards for our goals are available. Often there are even too many to choose from, so we rather propose guidelines to select an appropriate standard. Section 6 contains some agreed upon guidelines for selecting standards for multimedia and 3D.

4. Cultural heritage data

Cultural heritage data usually consists of different types of raw data (text as well as multimedia) and relationships between them. To keep all this related information together, we group it into a data item, which we'll call a "Cultural Heritage Data Object" (CHDO) from now on. An example of such a CHDO is a 3D representation of an object. The CHDO stores not only the 3D model itself, but also the way to obtain it (original scan data, information about processing steps and their parameters). Furthermore other annotations (e.g. in CIDOC-CRM format), other metadata (e.g. intellectual property rights) and links to other objects and descriptions can be provided.

Of course, a CHDO can not hold all the cultural heritage data of one application. For instance, all information on an archaeological site, ranging from 3D models to annotations and para-data, is too extensive to store in one object. For practical reasons CHDOs must have a finer granularity. As long as proper linking tools are provided, this is not a real limitation.

4.1. The diversity of cultural heritage data

In the different cultural heritage applications one finds quite a diversity of data characteristics. This is illustrated by the following three aspects: classes of typical usage, types of data, and the required accuracy. The CHDO should be able to accommodate all these different aspects.

Typical usage. The table below summarises the different requirements for the three different classes of usage: preservation, exchange, and presentation of the data.

	Preservation	Exchange	Presentation
Sustainability	important	less relevant	irrelevant
Standards used	few non-proprietary	few	application driven
Extend standard	avoid	allow	allow
Compactness	desirable	desirable	irrelevant
Fast running	irrelevant	less relevant	important
Fast development	irrelevant	less relevant	important
Wide use	irrelevant	desirable	desirable

An important aspect of cultural heritage data is its sustainability, but this is only important for preservation data, i.e., data which should be preserved for at least some decades. The life span of exchange data is much smaller and its format is more determined by the properties of the tools which are exchanging data. On the other hand, the life span of display or presentation data solely depends on the presentation application. Its storage format must be optimally tuned to the (speed) requirements of the application. Typical presentation standards, such as SMIL, are not very suited for preservation purposes.

The preservation format is usually considered to be the master format. This means that converters must be available to transform the preservation format into the exchange and the presentation format. Therefore, reusing the preservation format as exchange and/or presentation format reduces the conversion effort. Furthermore, one has to be aware of the fact that every conversion may lose information and/or accuracy. This is one of the reasons to avoid in-place processing of the preserved data.

Data types. Cultural heritage data usually has a heterogeneous nature. The tools work on different types of recorded data (images, 3D models, GIS information, text notes ...) as well as on different types of semantic data (meta-data and para-data, such as provenance, interpretation and other knowledge). For a peer review of the interpretation it is necessary to store all the information on the data processing steps, including programs and scripts used together with their parameters. Additionally, presentation applications often require user interaction, so they need to store animation and event processing.

A very important aspect of all these cultural heritage data is their relations, including the hierarchical aspects. Therefore the linking facilities must be rich enough to support all kinds of links, including links to a part of an object (e.g., the head of a 3D model of a statue).

Required accuracy. From the accuracy point of view, two different classes of data can be distinguished:

- Recorded data is recorded or generated data, which might be used later for analysis and interpretation. This data must be preserved with a maximal accuracy, i.e., the accuracy of the recording equipment.
- Illustrations (e.g. textures, movies, and sound) are only meant to look or feel or sound good. The accuracy is less important here as long as the perceptual quality is appropriate. This data can be preserved with a minimal accuracy corresponding to the required perceptual quality.

These classes have quite different requirements on the formats, e.g., on the allowed compression. Since recorded data must be kept at maximum resolution, compression of recorded data must be lossless. On the other hand, only the perceptual property of illustrations must be maintained so they can undergo a lossy compression.

4.2. Choosing a CHDO format

The CHDO is the basic unit for storing cultural heritage data. It should at least be defined for preservation purposes, although it can be useful for exchange of data also. It contains multiple types of data, ranging from recorded data to semantic information. It can be linked to other information (including other CHDOs and parts thereof) to form another CHDO at a higher hierarchical level.

If it is used for data preservation, the CHDO should be write-once. This means one can add information and links to it, but never change or remove information. Modifying a preservation CHDO should not be done in-place; it must result in a new CHDO.

If it is used for data exchange, the CHDO must accommodate multiple standards or exchange formats. If tools support a common exchange format, it is very inefficient to convert it to another intermediate format just for the sake of having only one exchange standard. Of course, it's recommended to have guidelines in order to keep the number of exchange formats as low as possible.

There are three ways to implement a CHDO:

- One can select one data type as the primary type, select a standard for it and extend the standard to accommodate other types of information. Let's call this the "*primary type implementation*". An example would be to choose 3D as the main data type, select X3D as the corresponding standard and extend X3D with additional fields to store other data, such as CIDOC-CRM data.
- One can choose an appropriate standard for each type of data and provide linking information between the elements. A CHDO element is a part of the CHDO data, which has a certain data type. The CHDO becomes a collection of loosely coupled information and the CHDO coherence is solely defined by the linking information, distributed over the elements. Let's call this the "*juxtaposition implementation*".
- One can choose an appropriate standard for each type of data and provide a tight coupling between the elements by using a container format. The CHDO is the container together with its elements. The container format contains all top level CHDO information as well as intra-CHDO linking information. Let's call this the "*container implementation*". Figure 4 illustrates the concept: each subtype has its own format and elements use the standard designed for them; the container holds the relations and all data not specific to one element. Note that this is only an illustration of the concept: not all types are present and elements may be used more than once (e.g., more than one 3D model).

The main disadvantage of the primary type implementation is that it abuses the primary type standard (X3D in the example above). It relies on the fact that this standard allows standard extensions (which is the case for X3D). The problem with extensions is that they are not part of the standard, which makes them unusable for preserving information. One could get round this problem by providing all

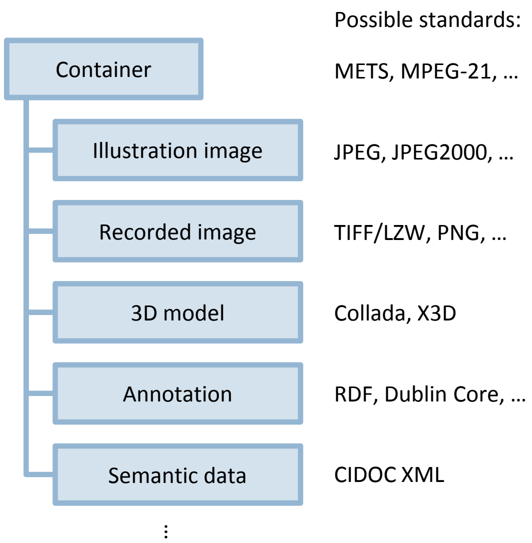


Figure 4: Using the container implementation to keep related data together.

the necessary tools to handle the extensions (e.g. an Epoch specific viewer) oneself, but this is very costly and very hard to sustain in the long run. A better solution is to have your extensions incorporated into the official standards by having them accepted by the international standard committees. Not only is this a very time consuming effort but it will only work if cultural heritage is a substantial part of the target group of the standard, which is not the case for X3D. So the primary type implementation can not be used as a preservation format.

The juxtaposition implementation solves this problem by using a specific standard for each type of data. However it still provides no solution for the relations between CHDO elements. Since these links are completely external to the CHDO, we not only need a fixed standardised linking procedure but each CHDO element must also have a registered identifier in order to guarantee sustainable links. Therefore the juxtaposition implementation is less suited as a preservation format.

Almost all of these problems are solved by using a container implementation: only extra-CHDO links have to be taken care of. But the container implementation has an important additional advantage: it decouples the standardisation aspects of all the different data types. With a container implementation we are not looking for a big global common infrastructure, but we divide it into elements, each with their own optimised implementation, and we look for a way to combine them into one CHDO ("divide et impera" as the Romans told us). This not only makes the problem manageable but it allows us to switch a standard implementation for one type without impact on the rest of the CHDO. This flexibility is certainly an advantage for exchange formats: we can replace JPEG by JPEG2000 without redefining the CHDO if all tools support it.

4.3. Requirements for a container class

Obviously, the formal design and choice of a container format is something that needs planning and studying. The general needs are:

- An existing standard should be preferred whenever possible. If that is not possible, extensions to existing standards or profiling the standard should be preferred over inventing something completely new.
- Care should be taken to choose the simplest solution that serves the requirements. Container classes can be very involved.
- Flexibility for accommodating future changes, because of unforeseen future requirements, should be possible.
- The standard should be believed to be stable and in use for a sufficiently long time.
- Intellectual property rights (IPR) issues should be incorporated. Not only for the whole data object, but also for each of its data items separately.

The technical needs are:

- Preferable it is based on XML. As it is W3C recommended and in general use, XML looks like the preferred way.
- Registration of unique identifiers to items or links should be possible. Think of e.g. DOI [DOI]. There are alternatives though.
- Immutable objects: data objects should probably be write-once. As they might be referenced from external sources, any change in them might invalidate these links. The only way to prevent this (apart from unpractical back-links), is not to allow changes to objects at all. Modifications to an object will create a new (version of the) object.

4.4. Existing standards for container classes

Quite some container class standards exist, but most of them are very much tuned to specific applications. Here we need containers suitable for "all multimedia" in the wide sense of the word. That restricts our choice to two standards specifically developed for this purpose, and one very general solution.

- *DIDL from MPEG-21* [DIDL]
  MPEG-21 standardises a multimedia framework. Of interest to the cultural heritage community, as far as is clear for the moment, is mainly Part 2, DIDL. As is obvious, the terminology is tuned for multimedia objects. Nevertheless, the possibilities of the framework are much wider.
- *XPackage and RDF* [RDF, XPKG]
  They can describe various resources and their associations. The major difference with DIDL is that RDF is less specific, less tuned to a particular application, but therefore also more flexible for non-standard needs.
- *METS* [METS]
  METS is a framework to structure relevant metadata (descriptive, administrative and structural) in the digital library world. It is increasingly used in digital libraries and museums.

At least each of MPEG-21/DIDL, RDF and METS can do the job. However, there is more active development in METS and more interaction between METS and other important standards in this field. So at this moment, we prefer METS as the import/export container format.

5. Cultural heritage information management

The previous section defined the CHDO to keep all related information together. However in practical implementations, it's impossible to store all data inside one CHDO. And even if it is possible, it may not be desirable for most applications, which access only small portions of the entire data set. So the actual granularity of the CHDOs in an application depends on the data usage as well as on implementation restrictions. This means that in most implementations we end up with a collection of CHDOs and relations between them. This is illustrated in figure 5.

For small applications one should be able to dispose of all overhead including the information management system. Applications which only manipulate semantic information will be implemented using dedicated software and database. It's clearly overkill to transform the semantic network into a network of CHDOs. This implies that the choice of a information management system should not dictate the formats and standards used to store the raw data, nor should it put restrictions on the tools needed to search or visualise the data.
But if one at a later stage wants to connect the semantic information to other information, it should be possible through an information management system. As illustrated in figure 5, this can be done by considering the semantic network as one data object or by splitting up the network in relevant parts, which can become part of other data objects.

In fact, all this is nothing else than the information management systems in the Digital Library approach. However, there are some specificities compared to a classical digital library system; rules on data items have to be much less rigid: new data can be added to old items, new types of data will appear, classification systems and interlinking can change during the lifetime of the data set and novel, unforeseen use of the data set cannot be excluded. Nevertheless, it would be nice if this approach could leverage on what has been developed in the digital library community before. Moreover, there have been ample discussions on long term aspects of storage, exchange and accessibility, which do need to be reassessed in the specific case of cultural heritage, but are largely identical. On top of that, there is existing software and freeware available, so implementation and maintenance costs can be minimised.

An extensive search for software that would implement a suitable modular content management system was done, and returned a set of tools that were usable and suitable for a test scenario, but none that was exactly tailored for Epoch-like environments. DSpace [DSP] is frequently used by libraries and archives and fully supports METS and MPEG-21/DIDL. Fedora [FED], not to be mixed up with the Fedora Linux distribution from RedHat, seems a bit more suitable for our purposes because of its 'dissemination objects' (kind of filter plug-ins in the server) and fine grained access control, but might have a smaller community than DSpace. Internally, recent versions of Fedora use FOXML (Fedora Object XML) for performance reasons. FOXML objects can be exchanged with other Fedora repositories, but import and export in METS is supported; METS1.4 and MPEG-21/DIDL plug-ins are under development.

6. Suggested standards and formats

Since Epoch is not a standardisation body, we provide only guidelines for selecting standards ("best practices") for intra-Epoch exchange (limited number of formats) and preservation in domains with Epoch interest. In our context, standards are about storing/exchanging information, not about visualisation.

Standards for audio, vector graphics, 2D images, and video were proposed, both lossless and lossy. More details can be found on the Epoch web site [MMS]. For 3D data the use of Collada is proposed. In Epoch a "Collada light" is used plus binary object representations. U3D seems very well suited for storing compressed 3D data.

Standards are frozen, so you have to switch to a new standard if

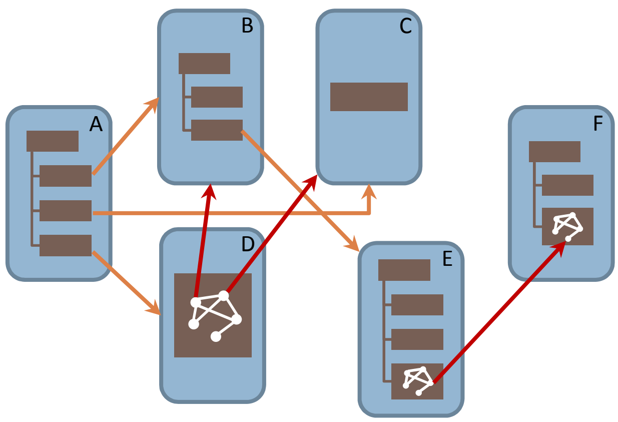


Figure 5: Linking data objects. The objects can be simple objects (C or the semantic networks D), or container objects. Links can represent an hierarchy (as from A to B, C and D) or direct relationships (between B and E). From inside an element links are not only possible to other objects (e.g., from D to B) but also to elements inside objects (e.g., from E to F).

you need new technology. Therefore, guidelines will evolve, certainly for exchange purposes.

7. Implementation examples

The infrastructure must be validated against real cultural heritage processes. This section describes two chains of tools which are not only used to demonstrate how to use the tools in an application, but also to test the interaction of the tools and the exchange of data between them. These tool chains were demonstrated at the exhibition part of Epoch's final 'Rome Event'.

A virtual visitor centre. This tool chain allows people to virtually visit sites which are normally closed to the public or which no longer exist. (The demo uses Pompeii as an example.) The visitor is guided through the virtual site by an avatar. At predefined places the visitor can examine objects (e.g. buildings or excavated objects) using a separate 3D-viewer. The visitor can also ask questions about the objects, preferably in his natural language.

Fedora is used as the content management system of this application. Different types of data are available about the 'House of Pansa', such as descriptions, multimedia data and metadata. This information is available on-line to the user (on a web site or on a mobile device), but a poster for the museum is also present (cf. figure 6). All presentations come from the same data source, but are tuned to the specificity of the users, their needs as well as their equipment hardware, with possibly minimal manual interventions needed from the person editing and composing the presentations.

Archaeological field survey. This tool chain addresses both ends of the cultural heritage chain of practice, from data collection to public presentation. During field walking, an archaeologist finds a small and interesting object. Information from MobiComp [MOBI] and 3D-scanning is combined with hand-written notes and stored in Fedora. Actors carrying a multimedia guide access the model when appropriate, according to their profile and preferences.

The infrastructure is based on Fedora and MobiComp. The models are stored in Fedora, while context elements are stored and handled by MobiComp. Contextual data in MobiComp is used to generate metadata for object models stored in Fedora. The models and metadata stored in Fedora are used by tools to generate context-dependent multimedia guides.

8. Conclusion

This paper shows that in spite of the diversity of cultural heritage applications some general concepts and guidelines can be established to provide interoperability of tools and sustainability of data. These concepts are not bound to the current state of available implementations, so they anticipate new technologies and solutions.

Acknowledgements

The research described in this paper were funded by the Epoch project and the Interdisciplinary Institute for Broadband Technology (IBBT). The concepts concepts described are the result of numerous discussions, by email as well as in meetings, between the partners of activity 3 of the Epoch project.

(a) Web format

(b) Poster format

Figure 6: Different presentation results starting from the same data of the House of Pansa in the Fedora repository.

References

[DIDL] J. BEKAERT, P. HOCHSTENBACH, H. VAN DE SOMPEL: Using MPEG-21 DIDL to Represent Complex Digital Objects in the Los Alamos National Laboratory Digital Library. *D-Lib Magazine* 9,11 (November 2003). http://www.dlib.org/dlib/november03/bekaert/11bekaert.html

[DOI] The DOI System. http://www.doi.org/

[DSP] DSpace. http://www.dspace.org/

[FED] Fedora Commons. http://www.fedora-commons.org/

[METS] Metadata Encoding and Transmission Standard. http://www.loc.gov/standards/mets/

[MMS] Multimedia formats in Epoch. http://partners.epoch-net.org/common_infrastructure/wiki/index.php/Multimedia_Proposal

[MOBI] MobiComp. http://www.mobicomp.org/

[RDF] Resource Description Framework. http://www.w3c.org/RDF/

[XPKG] XPackage. http://www.xpackage.org/

2 The EPOCH tools by topic

This chapter describes the tools developed in the Epoch project by topic. They were presented this way at the Epoch Conference on Open Digital Cultural Heritage Systems, held in Rome in February 2008. During this conference, the tools were also demonstrated to a wider public.

EPOCH Conference on Open Digital Cultural Heritage Systems (2008), pp. 1–7
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

UPGRADE and IMODELASER: Tools and Best Practice for 3D Data Acquisition and Processing for CH in Different Environments

S. Sotoodeh[1], D. Novàk[1], M. Sauerbier[1], D. Scaradozzi[2], A. Caiti[2], G. Conte[2], S. M. Zanoli[2], P. Drap[3], P. Chapman[4], A. M. Pascoal[5], L. Long[6], K. Hanke[7]
[1]IMODELASER, IGP, ETH Zurich, Switzerland
[2]UPGRADE, ISME, Università di Genova, Italy
[3]UPGRADE, CNRS, École d'Architecture de Marseille Liminy, France
[4]UPGRADE, Department of Computer Sciences, University of Hull, United Kingdom
[5]UPGRADE, Instituto Superior Téchnico, Lisbon, Portugal
[6]UPGRADE, Département des recherches archéologiques subaquatiques et sous-marines, Marseille, France
[7]UPGRADE, University of Innsbruck, Austria

Abstract

The acquisition and processing of 3D data for the documentation of cultural heritage objects is gaining more and more importance. 3D models do not only serve as a valuable means for visualization and presentation of objects, but also represent a form of digital preservation and can additionally serve as a basis for archaeological or architectural analysis. The two NEWTONs IMODELASER and UPGRADE deal with this task by developing practices and tools for efficient data gathering and 3D documentation. While IMODELASER aims to optimize terrestrial image-based 3D modelling by means of integrating terrestrial laser scan data, e.g. of buildings or other medium scale objects, UPGRADE concentrates on data gathering and representation of underwater sites.

1. Introduction

In EPOCH Work Package 3, after identification of missing tools in the processing pipeline, various "New Tools Needed" (NEWTONS) were defined to be developed and implemented. Two of these tools, UPGRADE and IMODELASER, deal with 3D data acquisition by means of combinations of different types of sensors and under different environmental conditions. IMODELASER concentrates on 3D modelling based on terrestrial images and laser scanning, aiming for an optimization of the obtained models by exploiting the strengths of both sensor types in order to overcome their specific weaknesses. These are mainly the weakness of automatic image matching techniques when dealing with homogeneously or repetitively textured surfaces and the inaccuracy of laser scan point clouds in terms of edge modelling. Therefore, in IMODELASER we aim for a combination of edges measured in images and highly dense point clouds for homogeneous surfaces yielded by laser scanning. As for IMODELASER in the terrestrial case, UPGRADE concentrates on 3D models but data gathering and representation are related to underwater sites. UPGRADE aimed for a widely automated generation of 3D models from underwater video, acoustic and navigation data. The output of the project was not only to automate the procedures of data gathering as far as possible, but also to improve the quality of the resulting, photo-realistic 3D models by combining the state of the art of underwater measurement instruments with suitable hardware and software. The problem concerns the design, test and development of a set of tools and best practices for collecting data from underwater archaeological sites by employing robotic vehicles and automatic devices under the supervision of archaeologists and engineers, possibly in cooperation, under some circumstances, with divers. The kind of data to get while navigating consists, essentially, of a set of georeferenced pictures in photogrammetric quality in order to construct a 3D model of the explored area in a virtual environment with a specified level of precision. This paper also presents how UPGRADE tools can be compliant with other NEWTON projects specifically presented in the terrestrial CH environment.

2. IMODELASER - Combining photogrammetry and terrestrial laser scanning

Within IMODELASER, we aimed for a widely automated generation of 3D models from terrestrial images and laser scan data. The goal was not only to automate the procedures of data processing as far as possible, but also to improve the quality of the resulting, photorealistically textured 3D models by combining the strengths of the two, in our case complementary, sensors deployed, namely digital still video cameras and terrestrial laser scanners, in order to overcome the specific weaknesses of each single sensor. The weakness of pure image-based 3D modelling consists of the fact, that automated image matching procedures, which are required to measure homologous points in two or more images to obtain accurate image orientations in a first step and 3D features which model the object in a second step, are error prone especially in homogeneous areas with low texture information. The consequences are no or wrong point measurements which then have to corrected manually. Contrarily, laser scanners have the capability to generate very dense 3D point clouds on such homogeneous areas, nevertheless they lack in terms of accuracy for 3D modelling of linear features, especially edges. Edges in turn can be extracted, partly automatically, from images with high accuracy. Therefore, IMODELASER aims for a procedure which uses the images for edge extraction, laser scanning to generate highly dense 3D point clouds and then, after automatic outlier removal, accurate co-registration of images and laser scan data and re-triangulation of the 3D features, determines a 3D model that represents the object highly accurate. For the demonstration of the capabilities of the developed and implemented algorithms, the ETH-Sternwarte, built in 1864 by Gottfried Semper in Zurich, was selected as a showcase object and modeled by means of our developed tools.

2.1. 3D surface reconstruction using laser point cloud and image data

3D surface reconstruction using point clouds from laser scanners is one of the practical approaches for surface modelling. It has to be conducted to generate a piecewise linear approximation of the object's surface. However some deficiencies of the data obtained by laser scanners , e.g. object boundaries, and the needs for automation of the modelling process make the potential for the employment of other sources of data.

If the sampling requirements are satisfied during scanning, point clouds obtained by laser scanners, including the geometry and the implicit topology, are used to reconstruct the object surfaces by triangular meshes [†]. Since the objects we are dealing with have a three-dimensional geometry and inherently the geometry of the point cloud, 3D triangulation of the point cloud is required. Although the 2.5D Delaunay triangulation gives a partial solution, it yields an incorrect topological surface model. Thus, surface reconstruction by a 3D Delaunay triangulation followed by surface extraction from the volumetric convex hull has to be performed. However, cleaning the point cloud from outliers is necessary prior to surface reconstruction

In this section we explain briefly a possibility to integrate edges from image measurements with point clouds from laser scanners to approach to a more complete surface model. The first subsection presents a method to clean data from errors and the second subsection explains the surface reconstruction and edge integration process briefly.

2.2. Outlier detection

Cleaning laser scanner point clouds from erroneous measurements (outliers) is one of the most time consuming tasks that has to be done before modelling. There are algorithms for outlier detection in different applications that provide automation to some extent but most of the algorithms either are not suited to be used in arbitrary 3 dimensional data sets or they deal only with single outliers or small scale clusters. Nevertheless dense point clouds measured by laser scanners may contain surface discontinuities, noise and diffrent local densities due to the object geometry and the distance of the object to the scanner; Consequently the scale of outliers may vary and they may appear as single or clusters. In [Sot07] we have proposed a clustering algorithm that approaches in two stages with the minimum user interaction and input parameters while it can cop with different scale outliers.

In the first stage the algorithm removes relatively large scale erroneous measurements and in the second phase it detects and removes the outliers that might not be as large as the first ones but according to the scanned object surfaces they are considered as wrong measurements. The algorithm has unconstrained behavior to the preliminary knowledge of the scanned scene and it does not suffer from the varying density of the points. The algorithm efficiency is assessed by a test on a simulated point cloud, which contained single and clustered outliers. The assessment is done with respect to a manual separation of outlier/inlier points. The type I and type II errors are estimated and the results are reported. In addition some examples in terrestrial laser scanner point clouds are presented and the behavior of the algorithm on the data sets are shown and discussed. Results show that the algorithm detects single and even clustered outliers almost without user interaction. Also, in case that the user editing is required, the result of the algorithm provides easier editing procedure due to the selection of point clusters rather than individual points. Some examples of the results are depicted in Figure 1.

2.3. Registration

Registration of images taken by CCD array cameras with respect to point clouds obtained by active sensors like laser scanners is necessary prior to the integration of the both data sets (texturing, overlaying, gap filling, etc). This is mostly done by the use of some targets visible in both data sets or by preliminary laboratory calibration, while the camera is mounted on the active sensor (e.g laser scanner). Nevertheless, there are various cases in which there are no common targets between the two data sets or the pre-calibration is not possible due to either using not mounted cameras or different acquisition times. Additionally, in case common targets are used, the mapping function that maps the point cloud to the image coordinate system is obtained over targets, thus the accuracy of the function is very much dependent on the distribution of the targets. More over, in case of pre-calibration, the calibration parameters may not remain stable if the camera has been dismounted for some movement purposes. Therefore a kind of on the job registration would be useful.

By orienting images with respect to an existing point cloud system it should be possible to see what improvements can be done when combining point cloud data obtained with different measurement techniques. We assume that the rough orientation parameters of the images are available in an arbitrary coordinate system. These parameters serve as approximations for an automatic process that incorporates a point cloud obtained by a laser scanner or a similar active sensor. By back projecting the points of the laser scanner point cloud into the images, the orientation parameters can be improved, by iteratively moving the two point clouds in a way that the distances between those two become minimal. As a result more precise orientation parameters are obtained, which in turn can be used to improve the existing point cloud with respect to edges and outliers that usually are present in point clouds from laser scanners or similar active sensor devices.

[†] Methods which use parametric sold shapes (plane, sphere, etc.) to perform object extraction are not referred here since this paper is about objects with arbitrary geometry

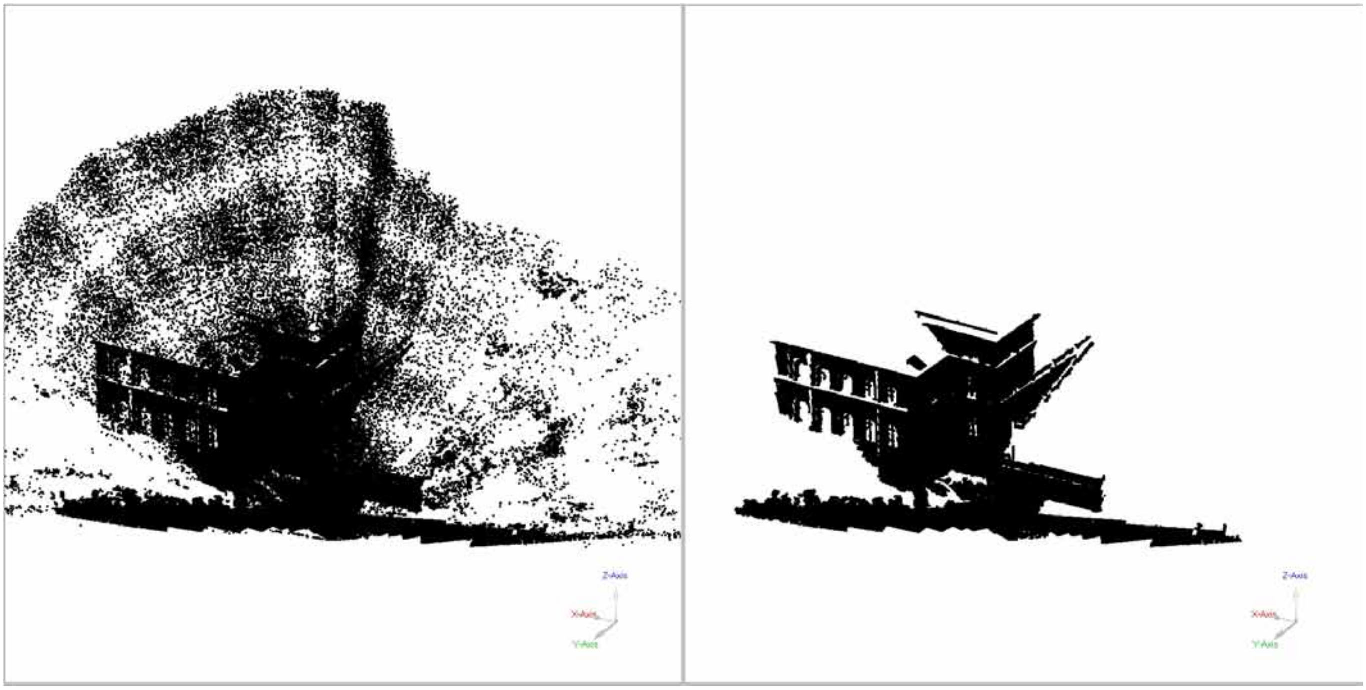


Figure 1: *Results of applying the proposed outlier detection algorithm on some scans of the Sternwarte building which are captured by Faro laser scanner. Left image shows raw point cloud and the right image shows the cleaned point clouds after applying the algorithm.*

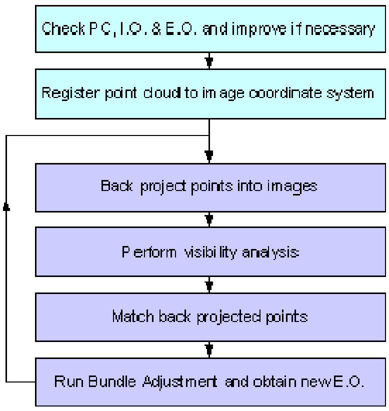


Figure 2: *Workflow of the registration procedure*

The data is divided into two data sets. One set is a dense point cloud that was obtained with a laser scanner. The other dataset consists of images taken of the object as well as measurements in the images. With the measurements an exterior orientation has been calculated with Photomodeler and as of such a sparse point cloud has been obtained. Figure 2 shows the suggested workflow of this task. The first two steps are considered pre-processing steps and are done manually. In a first step the point cloud has to be checked if gross outliers are present. If this is the case those have to be removed. Additionally the interior and exterior orientation parameters have to be checked and improved if they are not good enough. Then the laser scan point cloud has to be registered to the image coordinate system. The laser scan point cloud is rotated and scaled into the coordinate system of the photogrammetric point cloud. By doing so the photogrammetric point cloud remains fixed and the orientation parameters can be used for further steps. Afterwards the main processing chain is started, which mostly consists of automated modules:

- Backprojection of the laser points into the images
- Visibility analysis to reduce probability of mismatching
- Matching of the points

Two C++ tools are used which incorporate these three steps - the first tool is performing back projection and visibility analysis while the second one is doing the matching part. Two matching algorithms were implemented: cross-correlation and least squares matching [Gru]. Once the matching is finished, the matched points can be used as measurements for the bundle adjustment, which is performed with an external tool. After the bundle adjustment is completed successfully a new interior and exterior orientation of the images can be obtained. If the new orientation is acceptable the process can be stopped. If it is not then the points have to be projected back into the newly orientated images and the same procedures have to be performed again until the results are satisfactory.

The goal of this procedure is to have the images orientated precisely to the laser scan point cloud. By doing so the information of the images can be used to improve the model acquired by the laser scanner. Edges for example can be easily extracted from images. Since laser scanners do not perform well on edges the edge information can be transferred to the triangulation algorithm in order to have a better representation of edges in the final model. Another application is to perform outlier and blunder detection. By using the edge information it should possible to remove not only gross outliers but also blunders.

2.4. 3D Surface reconstruction

Having removed outliers, the point cloud is ready for surface reconstruction. First step is three dimensional Delaunay triangulation (Delaunay tetrahedralization). Unlike the two dimensional Delaunay triangulation which provides a surface mesh, the direct result of the Delaunay tetrahedralization (DT) is not a surface but a convex volume that surrounds the surface. Therefore the surface must be extracted using an additional procedure (3D surface reconstruction).

3D surface reconstruction from the 3D convex hull is an open problem even in computational geometry communities. Here, we have employed the most advanced approaches and we've contributed to the improvement of the approach using a neighborhood graph. The approach is based on the pruning of the initial surface model obtained by the extraction of the Gabriel graph from the 3D volumetric convex hull. The pruning is done in several steps based on topological relations, 2D manifold properties, a proximity graph, EMST (Euclidean Minimum Spanning Tree) and a neighborhood graph, KNN (K-Nearest neighborhood), followed by a hole filling process. Although the implemented approach is not a complete solution of the problem, it shows interesting results. One should notice that most of the process is done automatic and with the minimum user interaction. In addition no threshold or predefined value is required.

To get advantage of 3D edge information (from images) in the object reconstruction procedure, an edge-constraint 3D triangulation followed by an edge-constraint surface reconstruction algorithms are implemented. Since the surface reconstruction algorithm is based on the Delaunay properties of the prior 3D triangulation step, preserving these properties had to be considered. A constrained Delaunay tetrahedralization which has provably good boundary recovery is introduced by [Sch02], however a constrained three dimensional surface reconstruction algorithm is not well covered by the current literature. Following his algorithm, the surface reconstruction algorithm is modified so that it can accept the 3D edges as constraints in the process. The modification is done in both convex volume generation and the surface extraction process. The ultimate result is a surface that in the defined 3D edges it respected the 3D edges and avoid wrong interpolation in edge areas (see Fig. 3).

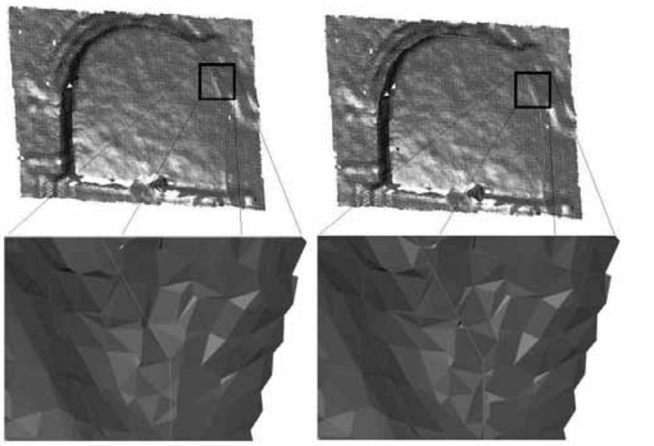


Figure 3: *Result of applying 3D edge constraint in the explained surface reconstruction process. Left images show surface without the constraint and the right images show the reconstructed surface after applying the 3D edge constraint (the line in the edge of the niche).*

2.5. Availability of the tools and conclusions from IMODELASER

The developed software consists of several standalone programs for the described processing steps. The input data on the one hand are 3D laser scan point clouds in ASCII XYZ format - other formats can be converted to this - and on the other hand oriented images - meaning that the coordinates X, Y, Z of the perspective centers of all involved images and the spatial rotation angles Ω, Φ, κ are known with respect to a common coordinate system - and measured 3D edges (OFF format). Currently, we investigate if parts of the software can be integrated into MeshLab (CNR Pisa), e.g. outlier detection and re-triangulation, to make them available also for further purposes. To make the developed workflow available, we opt to provide a service in terms of consulting and processing due to the facts that the photogrammetric processing requires expert knowledge from the data acquisition to photogrammetric processing and assessment of the results. Depending on the experience of potential users, different levels of support from our side can be offered. Regarding the developed method, a clear improvement of 3D edge modelling by means of triangulated surfaces could be demonstrated. The successful combination of image-based and range-based modelling opens new options for but also beyond cultural heritage documentation, e.g. when applied to aerial images in combination with airborne laser scanning. Though the developed tools partly still require long processing times, e.g. the registration of the images with respect to the laser scan point cloud, modelling quality can significantly be improved. Furthermore, there are options for speeding up the procedure by optimization of the programmed code. IMODELASER, or at least parts of the developed tool chain, can serve as one possible method for 3D data acquisition and modelling for other NEWTONs, e.g. 3DKIOSK and UPGRADE.

3. UPGRADE - Underwater photogrammetry and archaeological data enhancement

Over the last ten years, we have seen significant study in the development of efficient subsystems for underwater applications. This has been specifically driven by the offshore industry and European Commission in term of IT, research and new tools investment. Over the past decade, however, with the growing challenge of reaching deeper sites, human divers are being complemented with or replaced by Remotely Operated Vehicles (ROVs). These ROVs are equipped with sonar and vision equipment and, if required, tools to sample the seabed and manipulate artifacts. Regardless of how data are acquired, the crucial problem of overlaying the different data sets to generate composites of the interested site at both large and small scales has not yet been satisfactorily resolved. No off-the-shelf tools are available for these purposes. Several issues have not been satisfactorily addressed in the scientific literature and no efficient software tools are available to solve them. UPGRADE addresses an efficient way to equip a small class commercial robot in survey, for virtual reconstruction, of a geo-referenced portion of the sea bottom taking care of installation and synchronization of needed commercial and efficient subsystems. During the project, preliminary tools to equip the ISME's ROV for gathering, integration and fusing acoustic, optical and platform navigation data in surveying a prefixed area is presented. These tools should make the entire process largely automatic, and facilitate the construction of geo-referenced 3D maps of underwater areas. Another issue discussed in the paper is the automated generation of underwater images for photogrammetric purpose and for integration of Digital Terrain Models (DTM) at different resolutions, associated with the corresponding navigational data. This paper highlights the state of research developed in the EPOCH network in order to facilitate safer, faster, and far more efficient ways of performing archaeological site surveys and 3D reconstruction. In the following sections we describe the tools and strategies used in virtual reconstruction of an archaeological underwater site: the survey methods, the calibrating procedures, collecting photographs using ROV and divers and 3D data output formats.

3.1. Two different techniques for data capture

The photogrammetric survey is usually made by a set of photographs with the right overlap (around 60The photographs are taken 60If the site is not too deep (usually up to 60m), it is possible to survey both with divers (CNRS partner) or with ROV (ISME partner) [CZSG]. In this project our divers use a Nikon^{TM}D200 digital camera with a 14mm lens from Sigma^{TM}and two flashes Subtronic^{TM}. The digital camera was embedded in Subal^{TM}housing with a hemispherical glass, the ROV was equipped with Nikon DH2 digital camera, a 14 mm lens from Sigma^{TM}and two flashes Nikon^{TM}, SB800. The housing was made by COMEX with a flat glass and connector and the ROV interface was provided with ISME. Before survey the site is usually explored by multibeam (IST partner) in order to obtain a wide low resolution map of the sea bottom. The site is structured with scale bars (2m) and a set of markers (concrete blocks 15x15x10cm) or ropes are used in order to made a visible grid or path. During the survey the concrete markers or ropes are surveyed by the ROV, strip by strip, in order to cover the entire archaeological finds area [CZS*]. With the ROV, the photographs can be taken in two modes:

- Manually, the operator, looking at the scene through the lens using a small video camera installed on the photo camera.
- With a fixed frequency, decided according to the flash recharge time, ROV speed and altitude.

3.2. Multimedia calibration

Camera calibration in multimedia photogrammetry is a problem that has been around and documented for almost 50 years [BR, UPG80, UPG95]. The problem is not obvious, the light beam refraction through the different dioptres (water, glass, air) introduces a refraction error which is impossible to express as a function of the image plane coordinates alone [Maa95]. Therefore the deviation due to refraction is close to those produced by radial distortion even if radial distortion and refraction are two physical phenomena of different nature. For this reason we start to use standard photogrammetric calibration software and make a calibration of the set housing and digital camera. The distortion corrects in a large part the refraction perturbation. This approach is strongly dependent of the ultimate dioptre water/glass of the housing. To try to minimize the refraction error we can find on the market some housing with a hemispherical glass, which is the case of Subal^{TM}housing used with the diver. The other one, made by COMEX, the glass is a plate and the refraction action is much more important. In the second mission of the project (Tremiti Mission - Italy), special care is taken calibrating the panel: CNRS study and test two different panels in order to better compensate separately refraction and distortion (see Fig. 4).

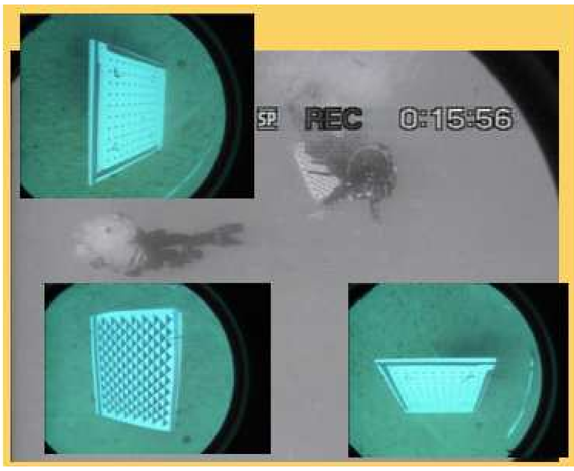


Figure 4: *Camera calibration procedure*

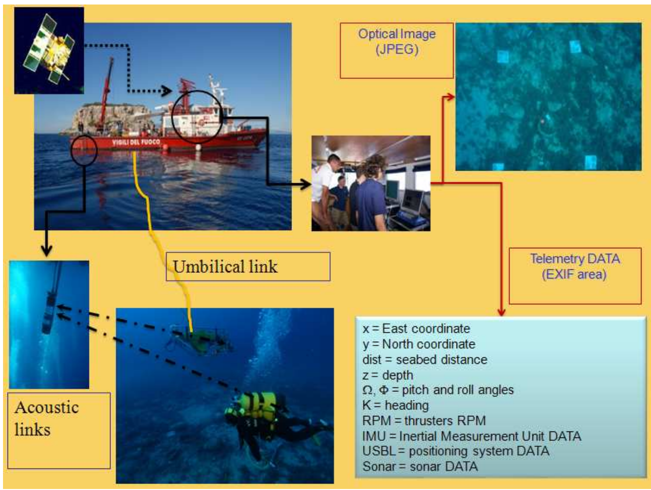


Figure 5: *Data capture system*

3.3. The tools for data capture

Based on the experience of the partners, during UPGRADE, two main systems are used to survey the sites: with an ROV or with divers. For ROV use, the main problem is to synchronize all data coming from different onboard sensors with the photograph command to the photo camera. One of the tools developed in the project is the EPL, a software library for generating georeferenced, enriched visual data from underwater photo and video. This software tool will generate EXIF postcards by associating acoustic, geographic and navigation data, acquired by the NGC system of the underwater robot, to photos and video frames. The software library can be included into NGC system of commercial ROVs and AUVs for JPEG/EXIF restitution of surveyed areas (see Fig. 5). During diver survey not all information can be synchronized, in particular the USBL system, which data are on the surface console, can be merged with information on the photo camera only after the end of the survey. A useful tool developed to enrich survey diverâĂŹs photograph is PUT, a small device to collect data during the flashing phase about photo camera orientation and acceleration. PUT (see Fig. 6) is a low cost device for underwater positioning and tracking of Divers.

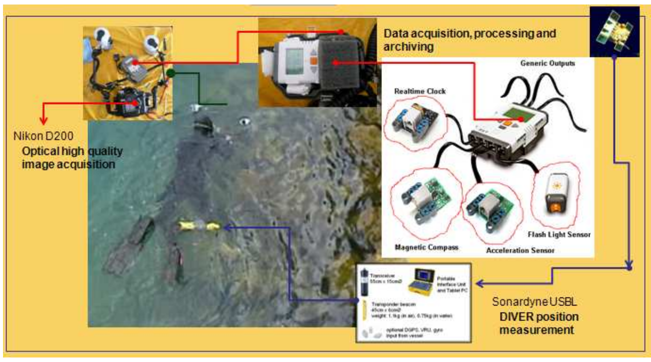


Figure 6: *PUT tool for data capture by divers*

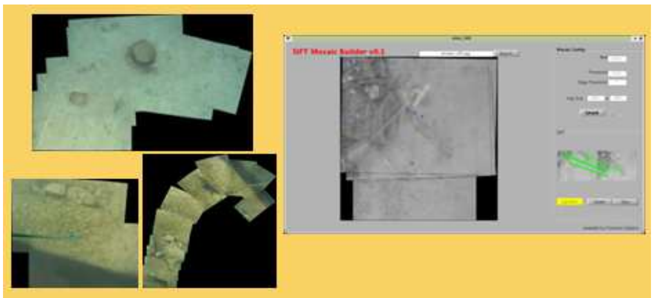


Figure 7: *SMuS tool for rapid mosaicing*

The device consists of a portable set of sensors, controlled by a microcomputer, for acquiring geographic and attitude data to complement photo and video surveys made by divers. Device construction is based on COTS and rapid prototyping methods. During the ROV survey or after the dive, with the data set coming from the tools developed, it is possible to view a partial reconstruction of the area with the photomosaicing tools SMuS. SMuS is a software suite for real-time photomosaicing from underwater photos and video frames of a surveyed area (see Fig. 7).

3.4. Site reconstruction

At the end of the survey, the data coming from the bathymetry mission, the photographic campaign and the archaeological knowledge of the artefacts are merged in virtual reconstruction of the 3D scene. All data can be stored in a relational database and a set of Java tools allows to wrap objects from the database and to produce a VRML or X3D representation. In the scene the seabed is textured using the oriented photographs. The textured seabed is obtained by triangulation of the points used to orient the photographs. UPGRADE has developed a tool to link each triangle to a set of possible photographs for texturing with the current used photograph mentioned. The result (3D points, triangle, oriented photographs) are written in XML with possible transformation to X3D and VRML. This way is very convenient to change the photograph used to texture a triangle or a zone. These data can be used to revisit the site for example with MVT. MVT provides marine cultural heritage professionals with the necessary tools to easily visualize and interact with bathymetric and other terrain datasets. It also provides pilots of Remotely Operated Vehicles (ROVs) to plan and visualize dives either offline or in real-time scenarios.

3.5. Conclusion

The UPGRADE project facilitates the construction of geo-referenced 3D maps of underwater archaeological areas. As in the other NEWTON projects, UPGRADE uses X3D and VRML representation but the data contained are different due to the environment characteristics and the instruments to measure the physical variables. At the end of the project we have the knowledge to relate some parts of underwater archaeology with what has been discovered and structured in other projects although it is not easy to have a common infrastructure between terrestrial and underwater data collection and representation. Because of the unique nature of the UPGRADE project, there is little overlap between the other NEWTON projects such as AMA, the context aware work of CIMAD, the avatars of CHARACTERISE and the AR visualization of 3DKIOSK. At the end of NEWTON and the EPOCH project, UPGRADE has provided a very specialised set of tools specific to the underwater archaeologist that also considers what the terrestrial scientist does and may have in mind for the future. A future opportunity, taking the advantage of new technology for underwater uses, is to join together both worlds starting from EPOCH results, best practice and tools.

4. Contact

IMODELASER:
Martin Sauerbier (martin.sauerbier@geod.baug.ethz.ch)

UPGRADE:
David Scaradozzi (d.scaradozzi@univpm.it)
Pierre Drap (Pierre.Drap@esil.univmed.fr)
Paul Chapman (Paul.Chapman@hull.ac.uk)

References

[BR] BASS G. F., ROSENCRANTZ D.: L'utilisation des submersibles pour les recherches et la cartographie photogrammétrique sous-marine. In *L'archéologie su-baquatique, une discipline naissante (1973)*. 5

[CZS*] CONTE G., ZANOLI S. M., SCARADOZZI D., GAMBELLA L., CAITI A.: Data gathering in underwater archaeology by means of a remotely operated vehicle. In *Proceedings, XXI CIPA international Symposium (2007)*, CIPA. 5

[CZSG] CONTE G., ZANOLI S. M., SCARADOZZI D., GAMBELLA L.: Robotics tools for underwater archaeology. In *ECSAC07 VII International Conference on Science, Arts and Culture. Science for Cultural Heritage. 2006*, ECSAC07. 5

[Gru] GRUEN A.: Least squares matching: a fundamental measurement algorithm. In *Close Range Photogrammetry and Machine Vision*, Atkinson K., (Ed.), Whittles, pp. 217–255. 3

[Maa95] MAAS H. G.: *New developments in Multimedia Photogrammetry*. Tech. rep., Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology, 1995. 5

[Sch02] SCHEWCHUK J. R.: Constrained delaunay tetrahedralizations and provably good boundary recovery. *Proceedings, 11th International Meshing Roundtable, Sandia National Laboratories* (2002), 193–204. 4

[Sot07] SOTOODEH S.: Hierarchical clustered outlier detection in laser scanner point clouds. *International Archive of Photogrammetry and Remote Sensing XXXVI*, 3/W52 (2007), 383–388. 2

[UPG80] *Manual of photogrammetry*. American Society of Photogrammetry, 1980. 5

[UPG95] *New developments in Multimedia Photogrammetry*. Wichmann Verlag, 1995, ch. Optical 3-D Measurement Techniques III. 5

EPOCH Conference on Open Digital Cultural Heritage Systems (2008), pp. 1–6
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

MeshLab and Arc3D: Photo-Reconstruction and Processing 3D meshes

P. Cignoni[1], M Corsini[1], M. Dellepiane[1], G. Ranzuglia[1], M. Vergauven[2], L. Van Gool[3] [2]

[1]Visual Computing Lab, ISTI - CNR, Pisa Italy
[2]K.U.Leuven ESAT-PSI [3] ETH Z§rich D-ITET-BIWI

Abstract

The paper presents a complete free software pipeline for the 3D digital acquisition of Cultural Heritage assets based on the use of standard photographic equipment. The presented solution makes use of two main tools: Arc3D and MeshLab. Arc3D is a web based reconstruction service that using computer vision techniques are used to vision technique based on the automatic matching of image features compute for each photo a depth map. MeshLab is a tool that allow to import and process these range maps in order to obtain a ready to use 3D model. Through the combined use of these two systems it is possible to digitally acquire CH artifacts and monuments in affordable way.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Line and Curve Generation

1. Introduction

Two partners of the Epoch Network of Excellence [EpochNOE], the ESAT-PSI lab of K.U.Leuven (Belgium) and the Visual Computing Lab of CNR-ISTI (Italy) have been active for a long time in the field of 3D reconstruction, albeit each in their own specific sub-area, passive reconstruction from images, 3D mesh processing and visualization, respectively. They decided to combine their strengths and set up a low-cost 3D reconstruction pipeline to be used in the cultural heritage field. The idea is that only a digital photocamera, a pc, and an Internet connection are necessary for a user to reconstruct scenes in 3D. The central part of this approach is a web accessible service, called Arc3D that allow to reconstruct raw 3d data directly from photo sequences, whose results are then processed by the system, here described, called MeshLab.

2. Arc3D Overview

Figure 1 shows a schematic overview of the client-server setup of the 3D webservice. The client- (or user-) part is located at the top, the server side at the bottom. On his PC, the user can run two programs, the upload tool and the modelviewer tool, indicated with A and B.

In the upload tool, images can be imported that were taken with a digital camera. Once authenticated, the user can transfer these images (C) over the Internet to the EPOCH server at ESAT. There, a fully automatic parallel process is launched which computes dense 3D information from the uploaded images. The parallel processes are run on a cluster of Linux PCÕs (D). When the server has finished processing, the user is notified by email and the results can be downloaded from theESAT server by FTP. These results consist of dense depthmaps for every image and the corresponding camera parameters (E).

A more detailed description of the complex inner working of this system, and how it *magically* succeed to reconstruct the basic 3D data from simple sequences of photos, can be found in [VVG05]; in this paper we mostly describe the second part of the whole pipeline the one that brings to raw reconstructed data to 3D objects. For accessing to the service, downloading the uploading software add trying the reconstruction service visit [Arc3D].

3. MeshLab general architecture

MeshLab was designed as a general 3D mesh processing system tool with the following primary objectives in mind:

- **ease of use**. The tool should be designed so that users without high 3D modeling skills could use it (at least for the most basic functionalities)
- **CH/3D scanning oriented**. The system should try to stay focused on mesh processing instead of mesh editing and mesh design where a number of fierce competitors already crowd the software arena (notably blender, 3D Max, Maya).
- **Efficiency.** 3D scanning meshes can easily be composed by millions of primitives, so the tool should be able to manage them.
- **Sustainability.** The project should be able to grow and self sustain itself for at least some years.

As a result MeshLab presents itself a *mesh viewer* application, where a 3D object, stored in a variety of formats can be loaded and interactively inspected in a easy way, by simply dragging and clicking on the mesh itself. MeshLab supports a ever growing variety of 3D formats (all the most commons format are supported) to accommodate the broadest set of users. Once loaded a mesh the user can work on it by mean of a large set of direct parametric filters, that performs unattended automatic task like smoothing, re-meshing or

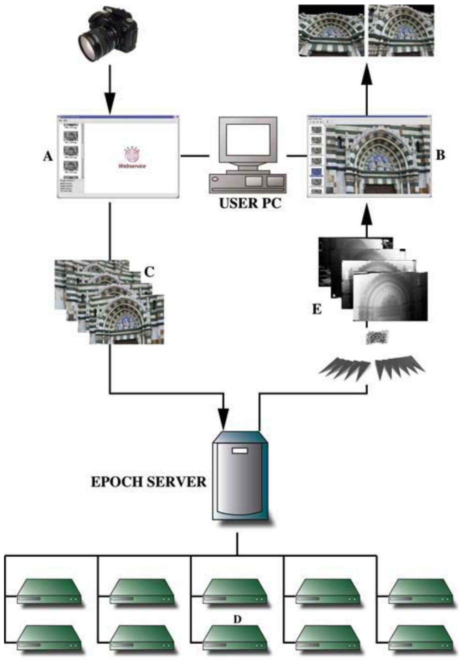


Figure 1: *Schematic overview of the client-server Architecture of the Arc3D web service. Images (C) are uploaded from the upload tool (A) on the user side to the server. There they are processed on a PC cluster (D). The results (E) can be downloaded via ftp and processed on the user PC with the MeshLab tool (B).*

simplifying, or by mean of interactive tools. Figure 2 shows an example of an interactive filter when the user drag the mouse over the mesh a local smoothing is performed on the touched portion of the mesh in real time; in this case the result is that the user is washing out some features of the object. No classical design-oriented feature are provided, structured editing of complex scene graphs is not supported by design. Multiple meshes can be loaded together and processed separately or together following a *flat* approach based on *layers*. The mesh processing functionalities that MeshLab currently provide are many and a short, incomplete, high level list of MeshLab characteristic is presented in the following:

- Interactive selection and deletion of portion of the mesh. Even for large models.
- Painting interface for selecting, smoothing and coloring meshes.
- Input/output in many formats:
  - import: PLY, STL, OFF, OBJ, 3DS, COLLADA, PTX, X3D, VRML
  - export: PLY, STL, OFF, OBJ, 3DS, COLLADA, X3D, VRML, DXF
  - Point Clouds support. 3D files that are composed only by points are well supported in PLY and OBJ format.
  - U3D support; MeshLab is the first open source tool to provide direct conversion of 3D meshes into the U3D format. with this format users can create, with just MeshLab and LaTeX, pdf files with 3D objects that can be seen with the current free version of Acrobat Reader.
- Mesh Cleaning Filters:
  - removal of duplicated, unreferenced vertices, null faces
  - removal of small isolated components
  - coherent normal unification and flipping
  - erasing of non manifold faces
  - automatic filling of holes
- Remeshing filters:
  - High quality edge collapse simplification (even with texture coords preservation)
  - Surface reconstruction from points (a ball pivoting variant)
  - Subdivision surfaces (loop and butterfly)
  - Feature preserving smoothing and fairing filters
- Various Colorization/Inspection filters
  - Gaussian and mean curvature
  - Border edges, geodesic distance, from borders
  - Non two-manifold edges and vertices
  - Self intersecting faces
  - Ambient Occlusion. An ambient occlusion field can be computed and stored per vertex
- Interactive Mesh Painting
  - Color Painting
  - Selection paint
  - Smoothing
- Measuring tool. You can take linear measures between points of the displayed meshes
- Slicing tool. A new tool that allows to export planar sections of a mesh in SVG format
- 3D Scanning tools
  - Alignment ICP based range map alignment tool, for putting meshes into the same reference space.
  - Merging of multiple meshes the Poisson surface reconstruction source code
- OpenGL Shader based rendering
- Large rendering. Images up to 16k x 16k pixels can be created for high quality poster printing
- The history of the all performed cleaning/editing actions can be saved and re-played on different meshes or saved for archival purposes.

3.1. MeshLab Development

One of the MeshLab ambitious goals was to create an open source application that is backed up by a sustainable development model. A common problem of many open source tool is that either they are small enough to be created and maintained by a small team, or, if they require a considerable development effort, the developer team need to be economically supported in some way. Many times the projects originates as research projects that are freed/opened (or simply released under some OSI approved license) at the end (or during) the (financially supported) research project. In these case the big incognito is the long term maintenance of the tool. In some cases the original developers, mostly for ethical and sentimental reasons, continue a low priority maintenance the software correcting small bugs and applying minor modifications, but great improvements are difficult to carry on. A direct consequence of this situation is that many many interesting projects lie in a rather abandoned state.

While we could base the development of the system on a large stable C++ library for mesh processing task [VCGLib], the whole

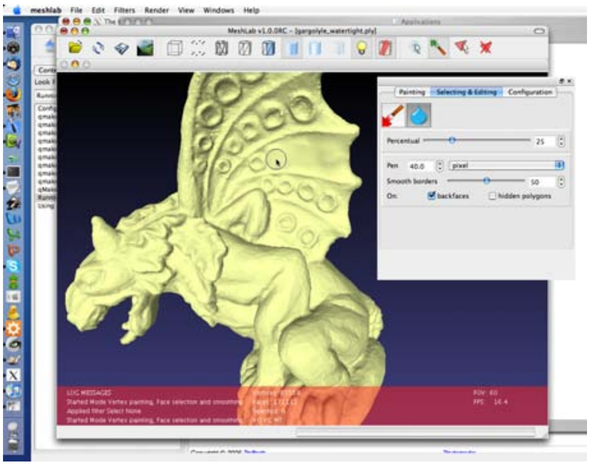


Figure 2: *A typical snapshot of MeshLab in action: the mesh under processing is interactively displayed and the user can work on it by mean of a large set of unattended parametric filters, or by mean of interactive tools, like the one shown in the above image, where the user is smoothing out some features of the object.*

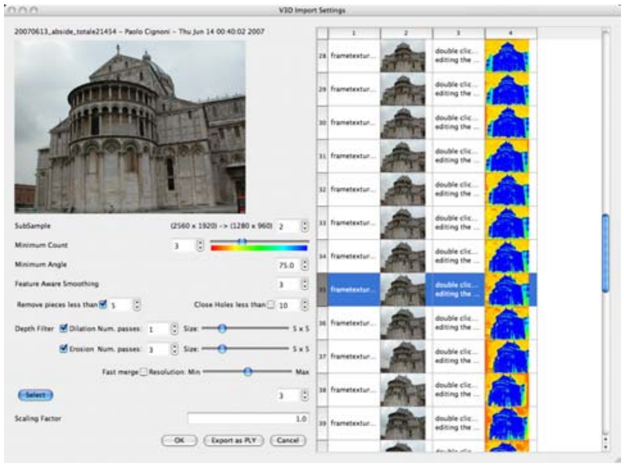


Figure 3: *The MeshLab interface of the importer for photo-reconstrcuted data from the Arc3D service. A synthetically color map depicting the quality of each range map is shown for each reconstructed photo. A number of option allow to fine tune the processing of the chosen range maps.*

task of building a consistent interface to present the library functionalities to the user (and add new missing pieces) was considerable.

Very large projects that can afford the support of a large community can rely on the loose sustain of many individuals, but large communities are difficult to be constructed. We decided for the above reasons for a very modular architecture and to subdivide the whole system into as small and as independent as possible pieces. These final pieces were small enough to be assigned, as course tasks, to students of FGT a Computer Graphics course at the Computer Science Faculty of the University of Pisa.

The idea proved successful, motivating the students more than classical 'mock-up' projects and fact that their code would have been actually used stimulated them to take their task with serious attention. On the other hand task that where assigned to low scoring students provided piece of code that not always was decided that they should be included in the core of the system.

As a result MeshLab started out in the autumn of 2005 an most of the 40k lines of code composing its core, have been written by willing students as part of their course work; each year students are adding new and new features to the system and they are actually making it to grow at a stead rate. The long list of many contributors to the MeshLab base code is well shown in the home page of the system.

4. MeshLab processing of Arc3D data

The output of the Arc3D web service is constituted by a set of range maps with texture and a per-pixel quality measure that denotes the number of found matches among all the images. For each range map all the intrinsic and extrinsic camera parameter are available, that makes easy to build, for each range map a simple point clouds with the points placed in the right spatial position.

Like all the data coming from digital acquisition devices they are affected in some measure by various kind of noise and errors. Knowing the kind of issues that are in the acquired data can help to remove or at least reduce them.

MeshLab provide a specific tool for importing the data produced by the Arc3D service. The tool is integrated inside meshlab as an input filter, when the author try to open a ".v3d" file created from the Arc3D the interface shown in Figure 3 is displayed to the user. All the images used for the reconstruction process are shown on the right with a color coded measure of the quality of the reconstruction process. The user can select some of the images (usually the best ones and /or the ones that cover in the better way the targeted object) and the 3D data generated from the Arc3D is used to build a set of range maps ready to be utterly cleaned inside MeshLab and then integrated in a single clean object. The user can also paint individual masks over the image for discarding portions of the images. A number of parameters are present in the left side to carefully drive the process of converting the raw 3D data into clean high quality range maps. Once the range maps have been imported a set of different operations have to be performed on them (cleaning, merging, aligning) in order to obtain a nice, ready to be used 3D model. In the next paragraphs we discuss some aspects of these processing steps

4.1. Weighted averaging and depth aware triangulation

First of all, some subsampling steps are usually necessary, current digital cameras has very resolute CCD an can easily generate tens of megapixel images. While this kind of definition is useful from the reconstruction point of view because it helps the robustness of the feature matching parts (at high definition even flat surfaces have some kind of texture), when converting these maps to 3D meshes this resolution is often not very useful. Infact the spatial accuracy of the reconstructed data is not uniform along the axis: along the xy image plane there is much more precision than along z depth direction. When you will integrate the various meshes the final quality will be affected by an error dependent from the maximal erorr of the single range-[maps/meshes so it is a good suggestion to subsample the images at the point that the error is more or less uniform along all the directions. Moreover the subsampling process is done

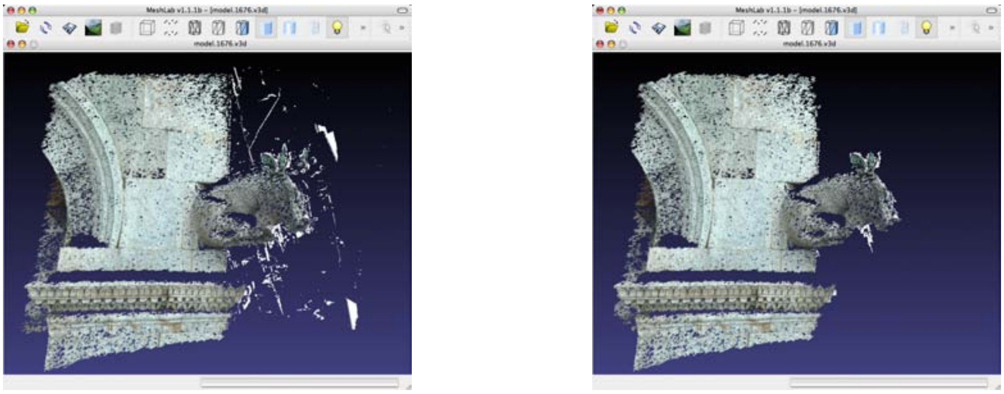


Figure 4: *MeshLab provides filters that automatically removes the small isolated pieces that float around each single range map. The same range maps is shown before (left) and after (right) the processing.*

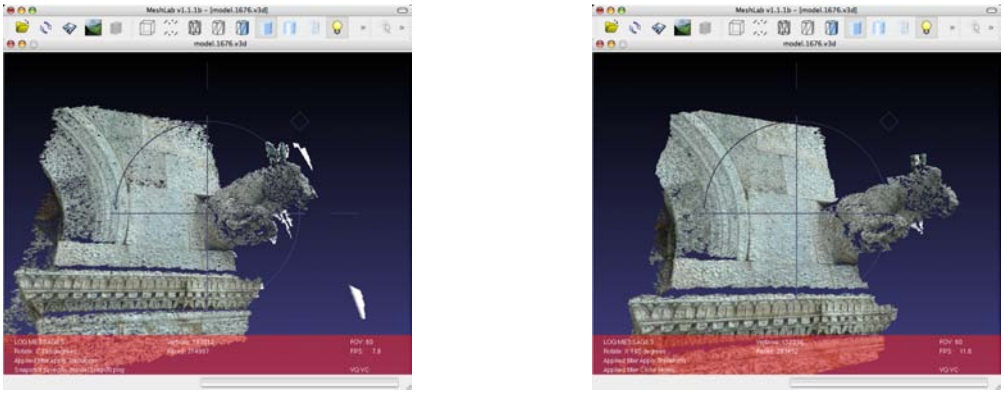


Figure 5: *The figure shows the combined application of erosion dilation approach for cleaning out the depth jumps outliers that sometimes appears. A set of small holes, that remains from outliers removal, is covered automatically by robust holefilling technique.*

4.2. Outlier management

From a practical point of view it is convenient to consider also procedures for the management of outliers. For a variety of reasons (wrong matching, non correct lighting setup, difficult surface reflectance properties) some of the acquired points can represent ghost surfaces that do not belong at all to the original surface. Luckily enough these points are usually rather isolated, or in other words they do forms small floating pieces far form the main surface. MeshLab provides a couple of automatic filters that allows to cope with these issues, these filters removes the all the small isolated connected components that are smaller than a chosen size expressed either spatially or in terms of primitives composing the component. Figure 4 shows an example of the filters that automatically removes the small isolated pieces that float around each single range map. The same range maps is shown before (left) and after (right) the processing.

4.3. Noise management

Image based acquisition methods suffers of noise problems that are in some way the opposite of the one that happens in traditional active scanning technologies. In fact the most difficult surfaces for passive techniques are the ones that are very flat, uniform and smooth. where the absence of prominent visible distinctive features, make difficult any matching procedure. The direct consequence of this behavior is that flat smooth surfaces can be affected by a considerable amount of noise. On the other hand it is possible to partially cure this problems by analyzing the original images used for the reconstruction process. In fact by detecting the portions of the image that do not posses enough features and marking them it is possible to apply them in a latter stage a heavier smoothing pass.

Moreover sometimes, typically when the photos are not taken in the perfect setting, the boundaries of high jumps region can present larger errors with eventually large and coherent outliers. Also these situations can be detected and corrected almost automatically by searching the boundary regions and applying a erosion-dialation approach that removes one or more strip of pixels from the dangerous jump zones.

Once you have removed all the outliers you have to choose if you should fill all the created holes and interpolate in some way the original mesh, This can be done inside meshlab, that provides some holefilling filters that can close all the small holes under a given size. Thy applies a variant of the algorithm presented in [Lie03], with various different heuristics for choosing the

Figure 5 shows the combined application of the techniques here described, large outliers and dangerous borders (look above and around the bull statue) are removed. A set of small holes (all over the wall), that remains from previous outliers removal, is covered automatically too.

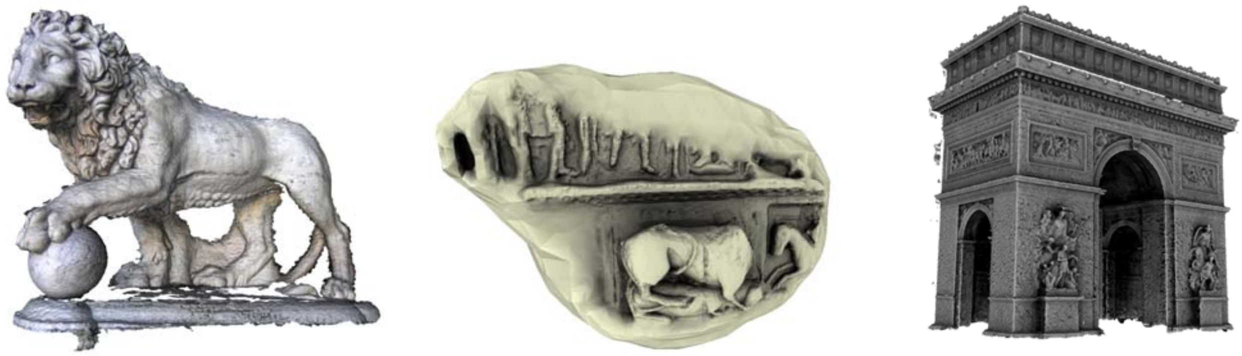


Figure 6: *Three samples of the meshes that can be obtained by using the Arc3D web service reconstruction and the MeshLab system. On the Left, a textured mesh representing one of the two lions in the Loggia della Signoria in Florence. On the center an untextured statue from the Portalada in Ripoll, a large portal in Ripoll, near Barcelona, on the right a reconstruction of the Paris arc du triomphe, untextured for better evaluation of the geometric quality of the geometric shape. The ambient occlusion lighting used to better enhance the shape features was computed with MeshLab.*

4.4. Merging

The individual range maps, after having been carefully cleaned and prepared, are ready to be integrated and fused in a single coherent model. All the range maps coming from a same sequence are already well aligned, so the next step is the one of to apply a surface reconstruction filter that fuses all the meshes in a single new mesh integrating the various parts together in a seamless way. MeshLab offers three different surface reconstruction algorithms. The first one is a interpolatory triangulation filter based on the Ball-Pivoting algorithm [BMR*99], that tries to build a surface connecting all the input points. This kind of algorithms do not work very well in presence of very noisy input data like the one coming from the Arc3D service. The other two are implicit surface approaches that uses the input data to build a implicit representation and then polygonalize it using marching cubes [LC87] variants. One is based on the Poisson surface reconstruction algorithm [KBH06], and it uses part of the original code provided by the authors themselves, and the second one, called *plymc*, developed at the Visual Computing Lab and described in [CCG*03], is an extension of the approach of [CL96]. The Poisson based surface reconstruction algorithm has the very nice characteristic that always build a watertight hole-free surface filling with an interpolatory surface all the missing parts, but currently do not support the color preservation. On the other hand the plymc approach preserves the color during the processing but leaves the unsampled areas as holes in the mesh; from a Cultural Heritage point of view this behavior could be considered a more *safe approach*. Figure 6 shows three samples of the reconstruction process. On the Left, a textured mesh representing one of the two lions in the Loggia della Signoria in Florence, the color was integrated from the many photos directly during the reconstruction. On the center an untextured statue from the Portalada in Ripoll, a large portal in Ripoll, near Barcelona reconstructed using the Poisson surface reconstruction algorithm; a watertight surface was robustly constructed, even if the input data contained only information on the front of the statue, building an interpolating surface even for the back of the model. On the right a reconstruction of the Paris arc du triomphe, done again with the plymc apprach untextured for better evaluation of the geometric quality of the geometric shape. The ambient occlusion lighting used to better enhance the shape features was computed with MeshLab.

4.5. Aligning

The aligning process it the step in which you take a set of different range maps of a same object, each one in its onw reference space and you rototranslate them in a single consistently aligned space. Usually in the traditional scanning pipeline [BR00], the aligning step, come before the merging, step but range map from each sequence get out already well aligned from the Arc3D service. The aligning issue arise when you wont to join two or more different sequences. In this case you have to align them together.

The Alignment process adopts a classical approach based on first, a pairwise local and then a global alignment [LR01, Pul99]. The initial placement placement of each mesh is driven by the user and require the selection of four or more points on both the meshes. This initial step is followed by an automatic process that fine tune the whole alignment . The alignment code used in meshlab is a derivation of the one used in Scanning Tools of the Visual Computing Lab [CCG*03], that has been used in a number of projects.

5. Conclusions and Future Works

The system has proved a success over the initial prevision. The last version has been downloaded 5000 times in the first three weeks, there are thousands of users from all the world with at least 500 users that have used it for opening more than one hundred of meshes, Users come from hundreds of universities and renowned commercial firms that have found MeshLab useful in contexts different from the original one of Cultural Heritages. Focusing on the CH context we would like to fill also another gap, the one of a broad distribution of the results of the 3D reconstruction of Arc3D. For this purpose we have designed a prototype, called PhotoCloud, to remotely browse the dataset generated by Arc3D, in a interactive, attractive way that do not try to offer the fine tuning possibility and high quality 3d meshes that you can obtain with meshlab tools but allow you to see all the original photographic information in the context of a rough 3D point-based model. A snapshot of the upcoming system is shown in figure 9

Acknowledgments We acknowledge the financial support of "EPOCH" (EU Network of Excellence, IST-2002-507382). A warm

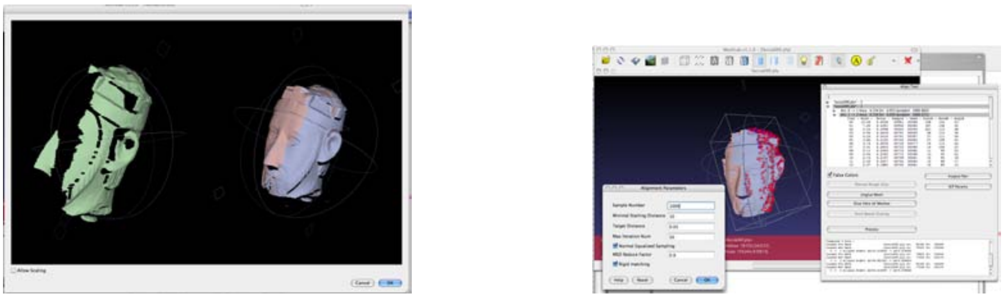


Figure 7: *A snapshot of the MeshLab Aligning tool, that allow to register different range maps, or in the case of data coming from the Arc3D web service, portions of a same object that have been reconstructed from different photo sequences.*

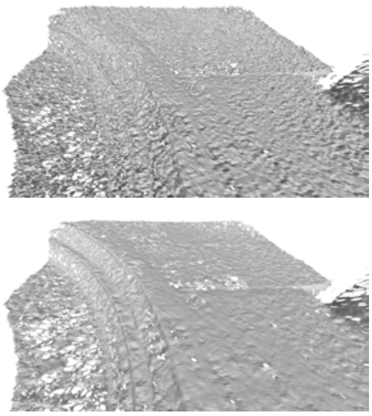


Figure 8: *Smoothing filters can help the process of noise removal. Sma*

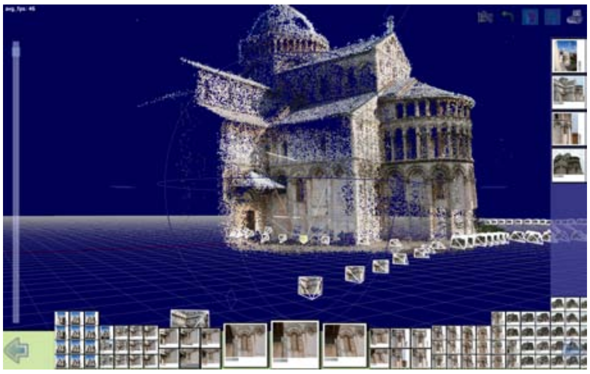


Figure 9: *A snapshot of the future remote viewer for photo-reconstructed models, PhotoCloud.*

thank you to the many developers and code contributors that have helped to build the current MeshLab system, in particular to M. Kazhdan and M. Bolitho for the Poisson surface reconstruction code and to A. Gfrei for the developing of the upcoming PhotoCloud.

References

[Arc3D] Arc3d: Automatic reconstruction conduit. More info on: `http://www.arc3d.be/`, 2007.

[BMR*99] BERNARDINI F., MITTLEMAN J., RUSHMEIER H., SILVA C., TAUBIN G.: The ball-pivoting algorithm for surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics 5*, 4 (Oct.-Dec. 1999), 349–359.

[BR00] BERNARDINI F., RUSHMEIER H. E.: 3D Model Acquisition. In *Eurographics 2000, State of the Art Reports Proceedings* (Aug. 24–25 2000), Eurographics Association, pp. 41–62.

[CCG*03] CALLIERI M., CIGNONI P., GANOVELLI F., MONTANI C., PINGI P., SCOPIGNO R.: VCLab's tools for 3D range data processing. In *VAST 2003* (Bighton, UK, Nov. 5-7 2003), D. Arnold A. C., Niccolucci F., (Eds.), Eurographics, pp. 13–22.

[CL96] CURLESS B., LEVOY M.: A volumetric method for building complex models from range images. In *Comp. Graph. Proc., Annual Conf. Series (SIGGRAPH 96)* (1996), ACM Press, pp. 303–312.

[EpochNOE] EPOCH: The European Network of Excellence on ICT Applications to Cultural Heritage (IST-2002-507382). More info on: `http://www.epoch-net.org/`, 2006.

[KBH06] KAZHDAN M., BOLITHO M., HOPPE H.: Poisson surface reconstruction. *Proceedings of the fourth Eurographics symposium on Geometry processing* (2006), 61–70.

[LC87] LORENSEN W. E., CLINE H. E.: Marching cubes: A high resolution 3D surface construction algorithm. In *ACM Computer Graphics (SIGGRAPH 87 Proceedings)* (1987), vol. 21, pp. 163–170.

[Lie03] LIEPA P.: Filling holes in meshes. In *EG/ACM Symposium on Geometry Processing (SGP)* (June 2003), Eurographics, pp. 200–206.

[LR01] LEVOY M., RUSINKIEWICZ S.: Efficient variants of the ICP algorithm. In *Third Int. Conf. on 3D Digital Imaging and Modeling (3DIM 2001)* (May 28th - June 1st 2001), IEEE Comp. Soc., pp. 145–152.

[Pul99] PULLI K.: Multiview registration for large datasets. In *Proc 2nd Int.l Conf. on 3D Digital Imaging and Modeling* (1999), IEEE, pp. 160–168.

[VCGLib] THE VISUAL COMPUTING LAB - ISTI - CNR: Visual computing library: a gpl c++ library for mesh processing. More info on: `http://vcg.isti.cnr.it/`, 2004.

[VVG05] VERGAUWEN M., VAN GOOL L.: Web-based 3D Reconstruction Service. *Machine Vision and Applications 17*, 6 (2005), 411–426.

EPOCH Conference on Open Digital Cultural Heritage Systems (2008)
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

The Presentation of Cultural Heritage Models in Epoch

Sven Havemann[1], Volker Settgast[1], Dieter Fellner[1,2], Geert Willems[3], Luc Van Gool[3], Gero Müller[4], Marting Schneider[4], Reinhard Klein[4]

[1]Institut für ComputerGraphik & WissensVisualisierung, TU Graz, Austria [2]TU Darmstadt & Fraunhofer IGD, Germany
[3]Center for Processing Speech and Images (ESAT-PSI/VISICS), K.U.Leuven, Belgium
[4]Universität Bonn, Institut für Informatik II, Germany

Abstract

The presentation of CH artefacts is technically demanding because it has to meet a variety of requirements: A plethora of file formats, compatibility with numerous application scenarios from powerwall to web-browser, sustainability and long-term availability, extensibility with respect to digital model representations, and last but not least a good usability. Instead of a monolithic application we propose a viewer architecture that builds upon a module concept and a scripting language. This permits to design with reasonable effort non-trivial interaction components for exploration and inspection of individual models as well as of complex 3D-scenes. Furthermore some specific CH-models will be discussed in more detail.

1. Introduction

Cultural objects in museum exhibitions are sometimes not easy to appreciate. They are too small, very detailed, behind glass, and too precious and delicate to let them be touched by visitors. It is particularly difficult to study and appreciate the amazing, intricate detail and the traces a long history has left on cultural artefacts.

3D technology can help tremendously to enhance the appreciation of museum objects. Our guiding vision is the idea of a *complementary exhibition*: Real artefacts in a museum exhibition are complemented by digital artefacts whose sole purpose is to deepen the understanding and appreciation of the real ones. A particular form of a digital artefact is the *digital replica*. To show it in a museum combines the authenticity of the real with the ease of manipulation of the digital. As a result, museum visitors become more engaged since they can actively participate. This approach is also quite successfully used in science and technology museums, whose number has greatly increased over the last years. School children are encouraged to actively acquire knowledge by trying out scientific experiments. – Cultural museums are unfortunately still lacking behind in exploiting this *edutainment* aspect.

It is important to emphasize that we do not recommend the use of technology under all circumstances. We advocate instead designing *modest systems*, where never the technology is in focus, but always the content. This implies, for instance, that technological gadgets must be extremely easy to use. They shall not require manuals or instructions, they must work robustly and, most importantly, they must behave in a predictable way. No bad surprises, no frustration, because that takes away visitors' attention from the artefact.

The same, however, applies to the CH professionals who preserve and present our common historical knowledge: We must offer them *enabling technology* that adapts to their needs and procedures, and not impose or presuppose a technologically biased view.

1.1. Strategic Vision

It is essential for Epoch that all the beautiful ideas and approaches as, e.g., shown on the VAST series of conference, find their way to the public audience through museums and exhibitions. In terms of interactive 3D presentations, much more is possible than what can be found today in average museums. We want to change this situation and make using interactive 3D a standard. The key features of our solution are therefore usability and affordability. The technical agenda of the Epoch project is based on the idea of a CH pipeline, a complete workflow that ranges from the archeologist that finds an artefact in the field to the presentation of the artefact in a museum. We think we should start to create a demand for content at the end of the pipeline in order to stimulate the development of a CH market based on Epoch technology on the whole pipeline.

1.2. A Concrete Target Scenario

A museum curator decides to make a new exhibition on the Etruscian period. He hires a company that digitizes the tiny golden brooches and small pieces of precious jewelry he is going to present. From a recent excavation campaign he purchases the scanned 3D datasets of different strata of the archeological site where the beautiful historic artefacts were found, as well as laser scans of the remains of houses, pillars, statues, etc. He puts all the digital artefacts into the 3D presentation software, chooses a theme (skin) that fits with the look of the physical exhibition, and uploads the new presentation to the 3D multimedia kiosks via network.

In the exhibition, visitors can see all the real artefacts, as usual behind glass. But in every exhibition hall there is also a 3D multimedia kiosk where the visitor can inspect the small beautiful artefacts interactively from all sides, which is not possible with the real. The visitor can also have a look at the archeological site where the artefact was found, which also brings up a few statues and other findings.

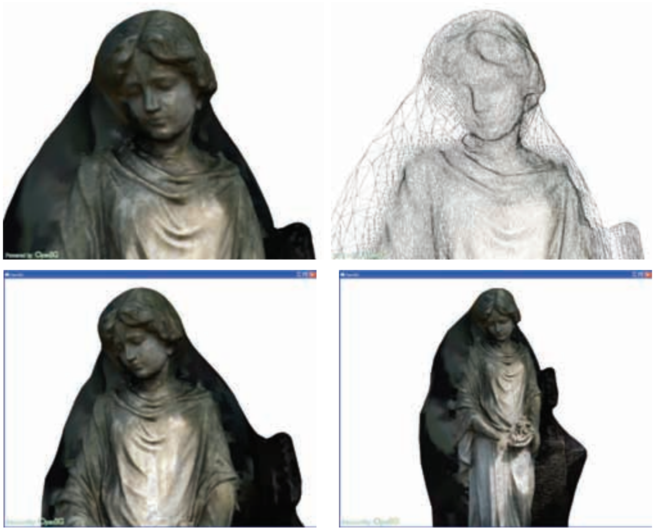


Figure 1: *Interactive inspection of a textured triangle mesh. The wireframe reveals how sparse the dataset actually is. A single* (r,g,b)*-texel per surface point gives realistic results only for very matte (Lambertian) surfaces, but not at all for complex, shiny materials.*

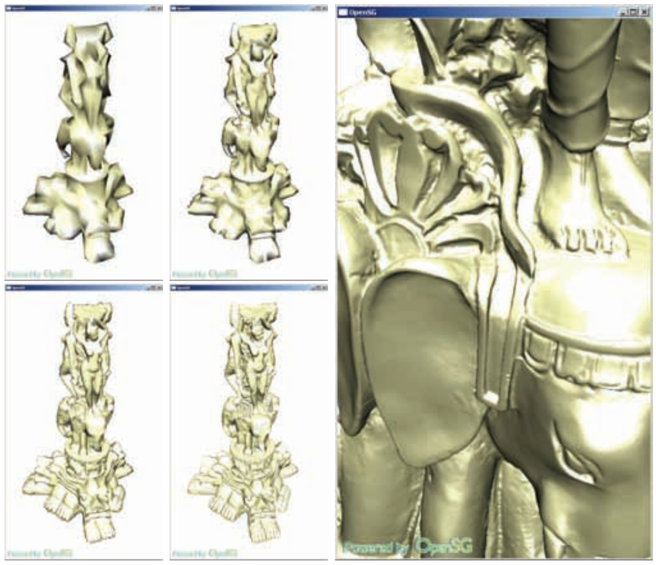


Figure 2: *An un-textured Nexus object. Nexus is a file format for massive multi-resolution meshes developed at CNR Pisa. It allows to first load very quickly only a rough surface approximation, which is then progressively refined during the interactive exploration to reveal finer surface details.*

1.3. The duality of Acquisition and Presentation

The purpose of this paper is to present some of the technical infrastructure that is necessary to let the described vision become reality. Creating an infrastructure is the focus of this paper, and also a great challenge: So far only a vast number of specific solutions to specific problems exist, usually described as research prototypes in scientific papers. It is very difficult to identify a common classification that applies to all of them. We follow a very rough schema using two classes to structure this paper:

- **Acquisition**: Produce sustainable digital replicae by measuring
- **Presentation**: Interactive exploration of a the acquired datasets

To support a specific shape representation means for the common infrastructure that solutions are operational for three things: for acquisition, for long-term storage, and for presentation.

The shape representation problem. The link between acquisition and presentation are obviously the datasets. The difficulty in designing a common infrastructure comes from the fact that there is a myriad of techniques to create and to store these data. Literally thousands of slightly different encodings (*file formats*) exist. But which of them to choose to support in the common infrastructure? Each has its strengths and weaknesses, each of them is being used within its particular community. For most of these communities, a rock-solid, sustainable long-term compatibility is not an issue, since this implies a development overhead that is not tolerable. But for CH as such, however, sustainability is crucial: Recorded historical evidence is supposed to be available forever.

The problem is illustrated in Figs. 1 and 2, which also allude to the complexity and the subtlety involved: Textured triangle meshes are probably the most-used data type for CH documentation today. On the one hand they are much too simplistic: Fig. 1 shows the incredible loss of information from the twelve 6 MP images (input data) to the sparsely sampled 3D mesh with vertex colors, i.e., only a single (r,g,b) triplet per mesh vertex (output). On the other hand triangle meshes are too complex: The size of the dataset in Fig. 2 is 526 MB, which is much too much keeping in mind that a typical display has only 2 million pixels. The obvious answer is to use multi-resolution meshes; but for this particular sub-problem alone hundreds of solutions exist, from Hoppe's Progressive Meshes to Cignoni's Block-maps.

1.4. Overview of the paper

The *Epoch common infrastructure* supports following shape representations with solutions for acquisition, storage, and presentation:

- **Synthetic reconstructions** created with standard 3D packages. The file formats of Max, Maya etc are for sustainability reasons not supported directly but only exported meshes (obj,ply)
- **Scanned 3D models** acquired by laser scanning or structured light are processed, e.g., using Meshlab to produce static (obj, ply) or multi-resolution models (Nexus, see Fig. 2)
- **Photogrammetric reconstructions** are obtained from sequences of photos that are converted to range maps by the Epoch Webservice [VG06] and then merged using Meshlab
- **Procedural cities** are synthetically generated using a shape grammar that encodes architectural rules [MVW*06] optionally taking into account archeological GIS data of building ground layouts.
- **Parametric shapes** are generated from a few high-level parameters by generative mesh modeling using the Generative Modeling Language [HF04, Hav05].
- **Scene graphs** contain a hierarchy (a tree) of nodes that are either coordinate transformations or individual objects (of the aforementioned types). Epoch uses the OpenSG scene graph, scenes can either be read from *Collada light* files [HSKF06] or from .osb files used for non-archival purposes.

New acquisition devices can have fundamental implications for also for the presentation, namely when new channels of information become available. Accordingly, the rest of the paper has two main concerns:

Acquisition: Models created using camera domes. This technology has a huge potential for the high-quality acquisition of CH artefacts with complex material on a mass-scale. However, including the produced models into the common infrastructure is also a technical challenge. To illustrate this, two recently developed acquisition domes will be discussed that both allow for the acquisition of geometry and appearance, but in slightly different ways.

Section 2 presents a single-camera dome for capturing the color and local surface orientation of objects by taking pictures where only the illumination direction varies. This allows not only for interactive relighting of CH models but also yields high-resolution surface mesh. This dome is also transportable to make on-site acquisition feasible.

Section 3 presents a multi-camera dome where the viewing and lighting directions can be varied independently. The captured data allow computing for each surface point a BTF, which is a much richer approximation to the true reflectance characteristic than standard texturing.

Presentation: The software infrastructure for the Epoch viewer. The Epoch viewer is not a stand-alone application, as the name might suggest, but rather a versatile toolkit that is supposed to be useful for all sorts of CH-related visualisation tasks. The underlying software infrastructure grew out of the 3D-Powerpoint tool developed in the 3D-KIOSK sub-project of Epoch (a so-called *Newton*).

Technically the Epoch viewer is a software component (under Windows an ActiveX control) showing only a 3D drawing canvas for interactive display of the OpenSG scene graph. This modest design makes it very flexible to use, since it can be embedded into a custom-tailored conventional GUI to give it the familiar appearance of a normal application. This way it can be used in a variety of ways, from scholarly studies (browsing of digital libraries) to public presentation in museums. Note that the latter has some features in common with Microsoft Powerpoint, which justifies the name *3D Powerpoint*:

- It has an authoring and a presentation mode.
- It is based on customizable presentation templates.
- It is made for non-programmers focusing on content.

This infrastructure is presented in more detail in sections 4 and 5.

2. The Camera Dome in Leuven, Belgium

Our acquisition system, first proposed in [WVM*05], allows for easy and cost-effective digitizing of small CH artefacts. The benefits of our system are the following:

- the hardware contains no moving parts and can be easily (dis)assembled, which makes it ideal for transporation and on-site acquisition.
- the many lightsources, combined with a robust photometric stereo algorithm, allow for the automatic and accurate extraction of the object's color, local surface orientation and depth map.
- the digital artefact can be studied by virtually relighting it, while several shaders are available to aid the scholar with the transcribation of digital cuneiform tablets.

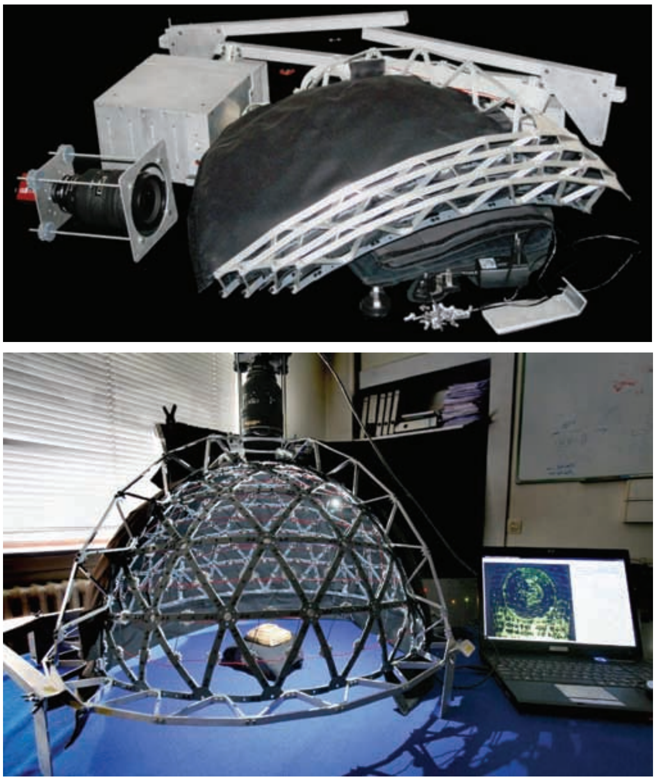


Figure 3: *The acquisition hardware. (top) Overview of all the separate components ready to put into a suitcase. (bottom) Typical acquisition setup.*

2.1. Motivation: CH documentation

Vast digitizing projects can be found throughout the last decade, and digital libraries are becoming increasingly mainstream. Besides the obvious reasons to digitally store and visualize artefacts, more and more digital models become the object of study for scholars as a surrogate for the actual artefact. The field of ancient Mesopotamian studies is a case in point.

The customary method to register and study cuneiform texts is a slow and often subjective process. The study requires a scholar with a specialization covering the specific contents, place of origin and time period of the text, and a scientific effort which can only take place at the location where the written document is kept because of conservation and political reasons. Cuneiform collections are spread all over the world, an additional disadvantage forcing scholars to travel around, which is both expensive and time-consuming. Identification or transcribation of the cuneiform signs is best achieved by taking the cuneiform tablet in one's hands to look for the best lighting which makes the signs legible.

To allow verification by future scholars, the cuneiform signs are interpreted by a specialist and an accurate line drawing of the text is produced. This is not only time-consuming, but correct transliteration and translation is thus completely dependant on the competence of the sole specialist who studied the specific text at a specific location.

2.2. Related Work

The *Cuneiform Digital Library Initiative*, a joint project between universities and museums, is one of the best known digitizing projects for cuneiform tablets and uses flatbed scanners to make 2D images and hand copies available on the internet. As a means of preservation, however, 2D images have inherent shortcomings as much of the three dimensional aspects of the clay tablets is not captured. Over the years, many solutions have been proposed to also capture 3D information by using laser and structured-light scanners.

Most recently, Hahn et al. [HBD07] have presented a hardware solution within the *iClay/Digital Hammurabi* project, which combines structured light with photometric stereo from a sparse set of images, using a DLP projector, a turn-table and a CMOS camera, to obtain a 2D+ and 3D model of tablets. Their set-up however requires moving parts and may therefore be unsuitable in some out-of-lab situations.

Mudge et al. [MVSL05] opted for a 2D+ representation of tablet using *Polynomial Texture Maps* (PTM) [MGW01] which approximate the luminance information at each pixel with a biquadratic polynomial. When a denser shape description was desired, the PTM data was registered with range data, again obtained via structured-light solutions.

Malzbender et al. [MWGA06], on the other hand, proposed a purely photometric stereo solution that recovers per-pixel estimates of surface orientation. While their solution has the great advantage of working in real-time, it does not account for specularities and shadows and is therefore not robust enough to extract high-quality shape information.

2.3. The Acquisition Process

The acquisition hardware consists of a upper-hemisphere of about 35 cm diameter on which 260 LEDs are uniformly spread out. A camera is positioned on the top looking down at the center of the hemisphere. The lights are being controlled via a custom built USB controller, while the camera is hooked up to the laptop via Firewire.

Furthermore, the dome was designed with portability and easy-of-use in mind. The hemisphere can be easily disassembled into 4 parts which, together with the supporting structure and USB controller, fit into a normal suitcase. The combined weight does not exceed 10 kg and build-up and tear-down take a mere 15 minutes. Combined with the camera and a powerful laptop, the whole system is therefore ideal to be used on-site. In figure 3, an overview of the disassembled components can be seen (top), together with a typical acquisition setup (bottom), ready for acquisition.

2.4. Processing Pipeline

Photometric stereo allows the estimation of local surface orientation and albedo by using several images of the same surface taken from the same viewpoint but under illumination coming from different directions. In order for a single side of a artefact to be digitized, it is therefore sufficient to position it in the center of the dome. After adjusting the lens and camera settings, a push of a button start the acquisition process which a takes a sequence of images, each time with another LED being lit.

While standard photometric stereo algorithms assume that the observed surface has a Lambertian behaviour, this is rarely the case. In reality, many artefacts have a reflective surface (e.g. coins) or a surface with a lot of depth variation (e.g. cuneiforms with sealings) that result in self-shadowing. Both cases render the Lambertian assumption invalid and force us to resort to a method which is robust against such complicating effects. An iterative approach has been proposed which computes the photometric stereo while selecting for each pixel those images for which the Lambertian assumption holds.

As we compute a normal at each pixel, a high-quality 3D surface can be obtained by calculating the depth map via multigrid-based normal map integration, as shown in Fig. 5. For more in-depth information, we refer the interested reader to [WVM*05, VWV06].

2.5. Digital Tablets

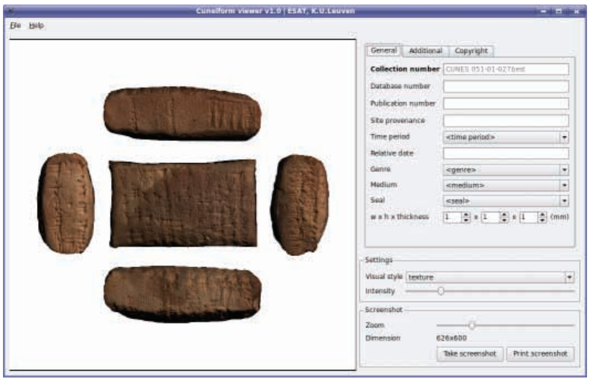


Figure 4: *Viewer application allowing for real-time rendering of virtual cuneiform tablets using many different shaders. Five sides of the unfolded tablet are shown in the left part of the split layout while the meta-data together with shader and print options are shown to the right.*

The virtual counterpart of a cuneiform tablet consists of a file containing all relevant meta-data together with a binary blob for each digitized side, which contains an albedo and normal map. As typically a high-resolution camera is used for acquisition, some compression is advised to keep the overall file size practical. The albedo maps are stored with 3 bytes per pixel using standard image compression techniques. The normal, on the other hand, are quantized and compressed from 3 floating point values into 2 bytes. Furthermore, all maps are cropped to their relevant regions, as selected by the user during acquisition. This results in a typical file size of 40MB for a virtual cuneiform tablet where all 6 sides of a tablet captured with a 5MP camera.

Once a cuneiform tablet has been digitized, the data can be loaded into a specialized viewer application, shown in figure 4, which allows the user to interact with the tablet in a similar way as if he was holding the actual artefact. Alternatively, a 3D polygonal model can be created from the extracted depth map. The availability of a tablet becomes therefore as easy as sending a file over the internet.

The tablet is typically shown as a unfolded cube, allowing the scholar to read'across edges'. The viewer allows the user to interactively shine a virtual light onto the virtual tablet with variable direction and intensity. By implementing the visualization purely on the GPU, all renderings are computed in real-time using the full resolution of the computed maps. Furthermore, several filters can be applied to enhance specific regions. Several of these filters are shown in figure 6. To the left, a typical relighting is shown which simulates a raking light over a tablet. The shaded version (middle) is obtained by showing the tablet without its color, which can improve rhe reading in some cases. At the right, the result from the line-drawing filter is shown which tries to separate the wedges from the tablet. This output can be used as a basis for a more objective hand copy as discussed in section 2.6.

Figure 5: *3D reconstruction of a tablet with prominent sealing. (left) obtained dense normal map from robust photometric stereo, (right) reconstructed 3D model.*

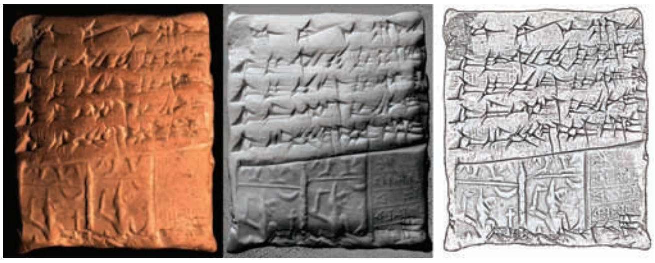


Figure 6: *Three different visualizations of the reverse of tablet CUNES 51-01-015. From left to right: virtual relighting without specularities, shaded version without albedo and extracted line drawing.*

2.6. Dissemination

The first trials of the hardware and software were run on the cuneiform collection of the Katholieke Universiteit Leuven. The benefit of this approach was that these texts were already published well in the past (OLP 3), the new recordings could therefore be evaluated alongside the result of this previous study. The recording, processing and application of the software were executed by the scholars of the Near Eastern Studies unit. Based on the feedback of these experiences, updates could be incorporated in the system. A second test phase was held at the Cornell University (NY, USA) on a never hitherto studied section of a voluminous collection. The Cornell collection reflects the typical complexity of a body of texts dating to various periods, coming from divergent provenances and encompassing different text genres. In this context, scholars of the Near Eastern Studies unit were asked to study those texts in the collection *CUNES* (see figure 6) dating to the late Old Babylonian period (ca. 1700-1600 BCE). To record the tablets, the hardware was transported in a standard suitcase and during a month all aspects of the system were thoroughly tested. The method of recording proved to be quite user-friendly for non technically educated researchers. Back in Belgium the texts could be studied successfully using the recorded results and by applying the digital filters in the viewer software.

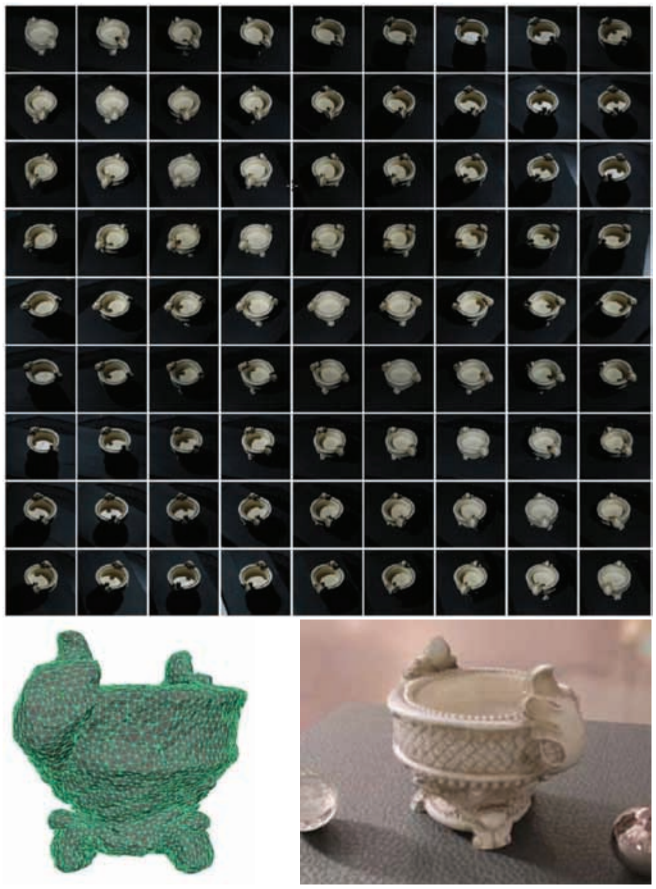


Figure 7: *The acquisition, reconstruction, and rendering of a CH artefact using the camera dome from Bonn. The workflow is explained in Fig. 9.*

The implementation of the system, especially on the Cornell University Cuneiform Collection, made several important benefits come to light. Instead of a long stay at the institution where the tablets are kept, only a short stay was now needed. The relighting application in the viewer did not only allow a clear identification of the signs, it also offers extremely good visibility of the sealings on the surface of the tablets. The study of these impressed sealings has always been difficult, but the combined use of the relighting system and the digital filters renders them clearly visible. One of the filters automatically extracts line drawings, which proved to be most useful for the correct interpretation of cuneiform signs and the sealings. Furthermore, this allows for a time-efficient and objective basis for the transliteration and translation process. To enhance the contrast between the surface and the impressed signs, one of the available filters extracts all color, which substantially improves the readability of the text. Before, this same effect could only be achieved by covering the surface of the tablet with a thin white coat of ammonium chloride (NH_4Cl), an intensive treatment now abolished.

The overall CH digitizing system has been mentioned in local Belgian written press and television and is featured in Euronews' Futuris episode *European Digital Library to contain all knowledge* [epi07]. From december 2007 to april 2008, the system is also being displayed in the national exhibition *"From Gilgamesj to Zenobia"* at the Royal Musea for Art and History (KMKG), Brussels, Belgium.

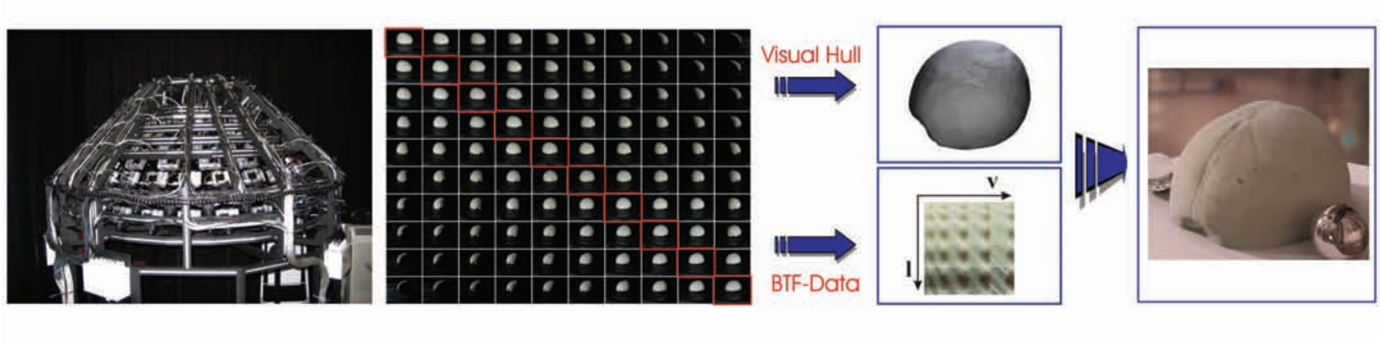


Figure 8: *The Acquisition Pipeline: The multi-camera array records 151 images per camera per light direction, resulting in 22801 images. Here, only the first ten cameras (from left to right) and the first ten light directions (from top to bottom) are shown. From the full set of these images, the BTF is constructed, while only a subset (the diagonal in this matrix notation) is used for the geometry reconstruction.*

3. The Camera Dome in Bonn, Germany

In order to produce visually appealing digital models of cultural heritage artefacts, a reconstruction of the 3D geometry alone is often not sufficient because accurate colour and reflectance information give essential clues for the object's material. Standard texturing methods implies severe limitations of material and lighting conditions. This in turn limits very much the usefulness of a digital replica, because moving, rotating, and seeing an artefact in different lighting conditions (environment maps) are indispensable for a decent understanding.

3.1. Camera Dome Approach in Bonn

With this situation in mind, a novel high fidelity acquisition system (see Fig. 8) to synchronously capture an object's 3D geometry *and* material properties in a very time-efficient and user-friendly way was developed in Bonn [MBK05, Ben07]. Our system exploits images from a multi-camera array to reconstruct an artefact's coarse to medium scale geometry using a GPU-based visual hull technique, resulting in a closed triangle mesh. In parallel, the images are also used to capture the object's appearance into so-called bidirectional texture functions (BTF) – a 6-dimensional texture representation introduced by Dana et al. [DNvGK97] which extends the common textures by dependence on light- and view-direction, and thereby allows for photo-realistic rendering of an objects micro- and mesostructure. The key contributions of our system are that it

- fully automatically acquires 3D-data, capturing an object's geometry *and* its visual appearance in form of bidirectional textures,
- faithfully reconstructs the object's mesostructure using BTF-techniques and therefore effectively overcomes the limited accuracy of the visual hull technique, and
- it is time efficient and very easy to use.

3.2. The Multi-Camera Grid

For fast acquisition of the images required to measure the BTF and to reconstruct the geometry of the object to be digitised, we use an array of 151 commodity digital still cameras mounted on a hemispherical gantry. By arranging the cameras into this array, the acquisition of the images required to construct the BTF textures is parallelised and no moving parts (e.g. a rotating stage) are needed. Therefore, the positions of the image sensors and the light sources can be calibrated in a preprocessing step which only has to be carried out if a camera has been replaced or after the whole setup has been transported. The low-level post-processing (geometric correction, colour correction) is fast enough to be done in parallel to the measurement.

3.3. Geometry Acquisition

In our current implementation, we first extract the objects silhouettes from the acquired 2D photographs by simple thresholding. As we use no backdrop in our setup, we set every pixel with a brightness of less then a certain threshold to be *outside*, the remaining pixel are *inside*. Together with the viewpoint information known from our setup, every outside-pixel in each image now defines a ray in scene space that is known *not* to intersect the object, whereas the inside-pixel define rays that intersect the surface at some unknown distance to the viewpoint. In the continuous case (pixel width $\rightarrow 0$) the union of all these intersecting rays would define a generalised cone that is guaranteed to contain the object. As this fact holds for all acquired images, the intersection of all these generalised cones (the *Visual Hull*, [Lau94, MBR*00]) describes a tight volume in space in which the complete object must lie.
The large number of acquired images and the (potential) need for finer grids make it impractical to actually traverse the grid following the *outside*-rays. Instead, we use a hardware-supported approach where we interpret the grid as a stack of binary 2D textures. The *inside/outside*-information is then efficiently collected by projecting every source image to each texture in the stack using projective texture mapping.

3.4. BTF Acquisition

The image-based approach to capturing the appearance of an object for later rendering is to take dense sets of images under controlled viewing and lighting conditions in order to sample its reflectance field appropriately. Since fidelity BTF measurements have huge memory requirements efficient compression techniques are needed. We use a clustered PCA compression [MMK04] [MSK06] which allows for an easy decompression and rendering using graphics hardware.
Statistical analysis like PCA, however, requires that data entries in the BTF are semantically correspondent – an assumption that holds for the raw data only under the assumptions of planarity, orthographic projection and directional light sources. This is not the case

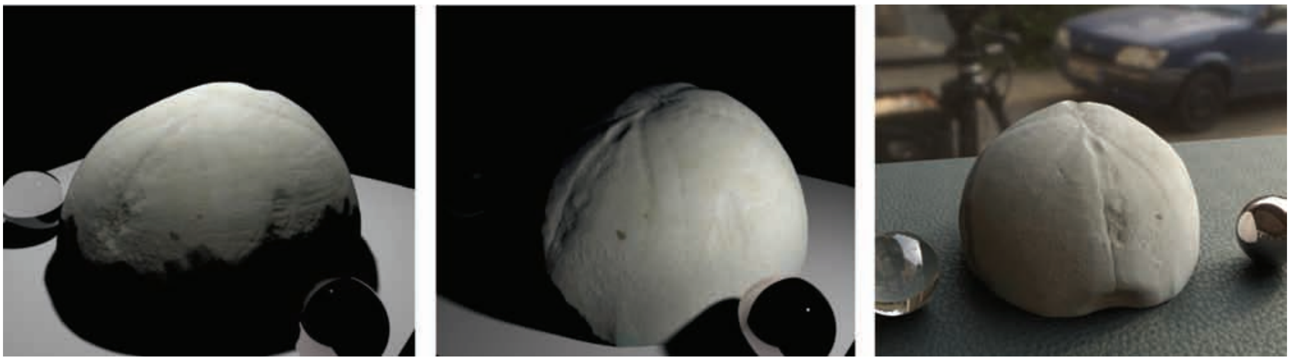


Figure 9: *Raytraced renderings of a captured and reconstructed echinite under novel lighting and viewing conditions. The left and middle image are rendered with a small area light source and demonstrate the fine geometric details captured in the BTF. The right image shows a relighting of the echinite with a complex image-based lighting environment captured in front of the Pädagogische Fakultät in Bonn. The ground floor in the right image is covered with a synthetic leather BTF courtesy of DaimlerChrysler AG.*

for our system since the dimensions of our acquisition setup cannot be considered "large" compared to the dimensions of the objects to be digitised. Therefore, we perform resampling of the raw BTF data based on a planar parameterisation of the reconstructed triangle mesh computed with the method of Degener et al. [DMK03] *before compression*.
Measuring BTFs generally consists of recording for every point on the surface its reflectance from each view direction under each light direction. For non-flat surfaces, however, the reflectance for some light and viewing directions will be zero (or close to) simply because of occlusion and/or self-shadowing. Using standard approaches, this *missing* information would be misinterpreted as a property of the material. Therefore, we first identify from all the points on the object surface those points that have incomplete measurements and perform a clustered PCA for them. The missing values of the remaining points are thern completed in a way that the overall reconstruction error of them is minimized.

3.5. Results

Our approach to reconstruct the geometry from the acquired images using visual hulls computed on the GPU is reliable and fast. Of course, identifying a nonconvex object using a silhouette-based approach inherently and inevitably implies neglecting some features of its surface geometry. Despite this general seemingly inaptness of the visual hull reconstruction, we were able to produce realistic images of our captured objects (see Fig. 9) because the neglected surface features are well-captured in their appearance using the BTF texturing techniques. To further improve the geometric reconstruction of the object, methods presented in [HS03, CGS06] can naturally be incorporated into our system.

4. Presentation: The Epoch Viewer

The presentation of 3D objects to a public audience is often considered a solved problem since numerous possible approaches exist: 3D standard tools, professional 3D presentation software, game engines, scene graph engines, and certain academic approaches. Usually the first idea would be to use 3D standard tools such as a VRML/X3D viewer, 3D embedded in pdf (Acrobat3D), Shockwave/Flash-3D, etc. Also for 3D modeling tools such as Maya, 3DStudio, Blender

Figure 10: *Models from the Herz-Jesu Church in Graz each reconstructed from 6-12 photos (6 MP) using the Epoch Webservice (KU Leuven) and the Meshlab software (CNR Pisa). Bottom: Water-tight models from the Poisson reconstruction filter in Meshlab, result of ambient occlusion shader, then multiplied with the acquired object color (Homer); and this with specular GLSL shader with per vertex color and per-pixel-lighting under OpenSG (Monster).*

presentation plugins exist. But these “closed” solutions can be immediately ruled out since we target location-based presentations with high-quality CH content, and smooth interaction with non-standard input devices. We could certainly program extensions to, say, a standard X3D viewer, but then we would be tied to this particular piece of software over which we have no control.

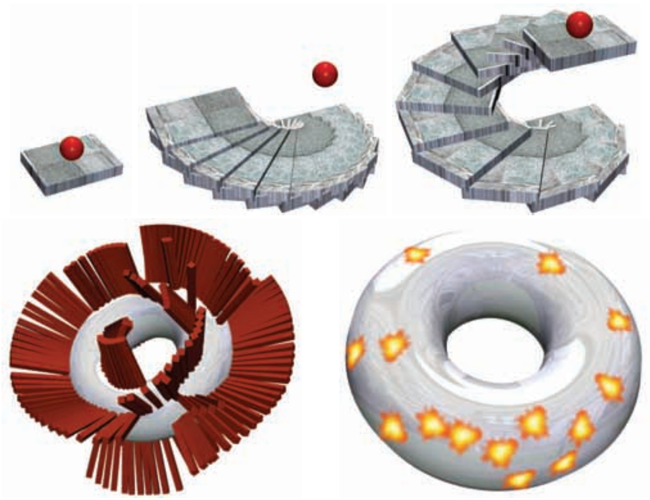


Figure 11: *Advanced interaction using GML scene graph scripting. Top row: Dragging the ball upwards unrolls the spiral staircase. Bottom row: The scene graph allows for very easy object distribution and orientation, for instance by using a coordinate frame computed from the surface normal at the picked position.*

A much better option would be to use professional 3D presentation software such as Virtools, Quest3D, Shark3D, OfficeFX, and others. These tools provide professional rendering quality, support a wide range of input formats, hardware setups, and all possible input devices, and they have impressive feature lists, ranging from physics engines over audio to networking. However, an in-depth evaluation some years ago of a similar tool, Realimation, revealed some fundamental problems with such packages. They are

- **monolithic:**
  Not a component, but a complete stand-alone application
- **proprietary:**
  Vital features may change from one version to the next
- **not low-level extensible:**
  They impose strict limits on what developers can access
- **not a modeler:**
  Every non-trivial piece of geometry must be imported

It must be noted, however, that these tools provide a remarkable degree of usability: their authoring environments are extremely interesting, also from an academic point of view.

For most of the professional game engines – the 3D engines database on www.devmaster.net/engines currently lists 279 of them – basically the same considerations apply: Game engines such as Torque, 3DGameStudio, Ogre, or Irrlicht are optimized for efficiency and use the latest graphics hardware effects. This matches also the expectations of museum visitors, as more and more people are becoming acquainted with game technology. The downside, however, is that content creation for games requires much low-level realtime know-how, much programming, has a low long-time sustainability, and for serious non-game applications requires extra effort to get rid of the game overhead.

A much more acceptable alternative is to use a "neutral" scene graph engine such as Coin/OpenInventor, OpenScenegraph, OpenSG and the like. They combine efficiency and support of (almost) latest effects with openness and extensibility. In fact, they are usually *designed for extensibility*. This is a huge advantage if, like in our case, we have to custom-tailor a 3D solution to the demands and existing standards of a particular application area, in our case cultural heritage.

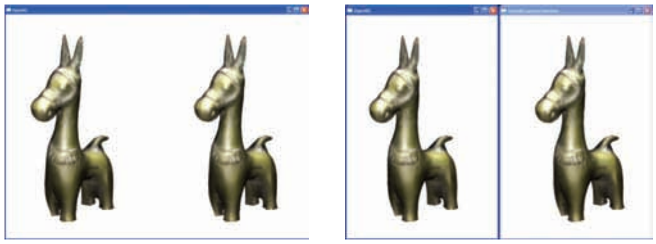


Figure 12: *Stereoscopic display modes, and BTF materials. In **horizontal span** mode the desktop goes over two physical screens, whereas with **dual view** a fullscreen application can cover at most one single screen. When both channels are overlaid using a pair of projectors and a decent per-eye channel separation, a stereoscopic effect can be achieved. – The object shown has a BTF material that contains more than one texel per point. It allows to approximate the reflectance function (BRDF) at each surface point much better. The BTF was measured using the camera dome from Bonn (cf. Fig. 7).*

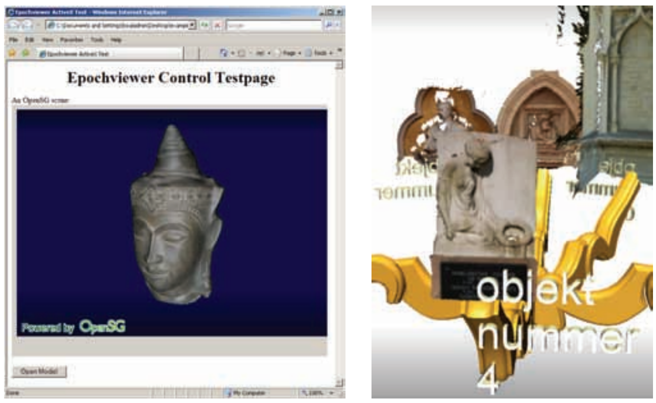


Figure 13: *A whole range of display scenarios, from quick browsing of a model database using a web browser to stereoscoping high-end rendering in a kiosk application, are possible using the Epoch viewer with its module concept.*

There are only two academic approaches for 3D presentation environment for CH we have found. First, Costagliola et al. [CMFP02] publish a configurable presentation environment particularly for guided tours. The other work is the *Virtual Inspector* tool from our colleagues from Pisa, Italy [CPCS08, CPCS06], which focuses on the display of huge meshes.

Our approach differs from those in that our tool provides an event system that allows scripting of all sorts of interactive behaviours, operating on a scene graph that is assembled at runtime. OpenSG, the underlying open source scene graph is extensible, so that also the rendering of huge meshes is possible with a suitable render node.

However, through this research we have come to the following catalogue of criteria for our envisaged solution.

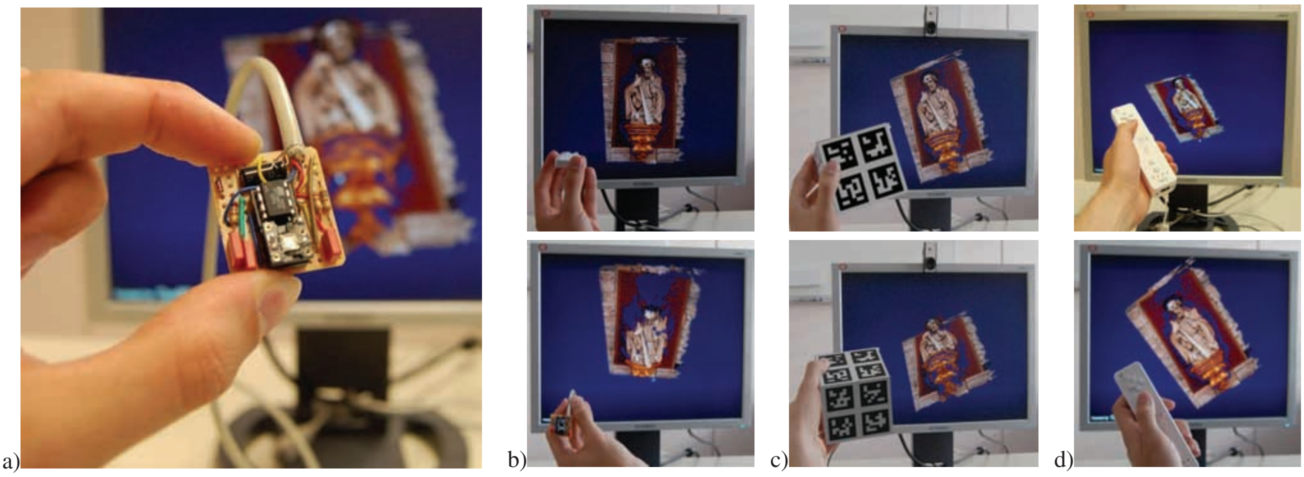


Figure 14: *Tangible input devices: (a,b) Accelerometer, can measure 2 of the 6 DOF of a static pose, as it can detect in which direction gravity points. (c) Array of* 2×2 *ARToolkit markers, to make the camera-based detection more robust agains occlusion by fingers. (d) Nintendo Wii controller, with accelerometer and integrated low-resolution camera, can in principle determine a 6-DOF pose plus 2-D acceleration.*

5. Feature Wish List

Drag-and-Drop Authoring: Every presentation consists of *layout* and *content*. Both are easier to generate and arrange in 2D than in 3D. As we want non-3D-experts to design digital exhibitions, the best solution would be to have *pre-defined layouts* that can be *filled with content* from the local file system or from the internet via drag-and-drop.

Versatile Hardware Setups: A viewer should support a wide range of input and output devices (see Fig. 14). To support basicallz any human input device is possible using the concept of *channels*; using a well-defined communication protocol and socket communication, input devices can connect via network. The output options range from a 3D window in a web browser (see Fig. 13) to cluster rendering in a CAVE. Also stereoscopy is a desirable feature, but many flavors exist (see Fig. 12).

Easy 3D navigation: Once the presentation is defined it shall run in a public museum. Average museum visitors shall be able to use the 3D kiosk systems without having to read instructions. This is demanding in particular for the notorious problem of 3D navigation: Users shall never get *lost in 3D*, not reach impossible view points or see nothing on the screen, nor get locked somewhere. We want to allow as much 3D control as possible and restrict it only as much as is necessary to enforce consistency.

CH Integration: 3D presentation is only the very end of a long processing chain. Interoperability requires standards. The presentation environment should permit to make use of any additional information attached to cultural objects present, e.g., in the Collada 3D format used in Epoch. In the long run, even using CIDOC/CRM should be an option, as pointed out by Havemann et al. in [HSKF06]: Every cultural artefact is part of a network of semantic information. The ultimate form of a *Cultural Presentation Browser* would be a tool that allows average museum visitors to navigate through this semantic network.

CH Sustainability: 3D presentations will be a new form of collecting knowledge about history and culture. Ideally, our knowledge about CH should be as long-lasting as the artefacts we show. This issue actually causes fierce reservations against the use of digital technology in the CH community. However, there is a tradeoff between using the latest technology and being independent from particular software and hardware versions. The solution we envisage is to use advanced, but *well-documented* file formats and algorithms. This way presentations can use state of the art hardware shaders and mesh compression, but are not deemed to be obsolete in five years.

Low-Level extensible: Today the usual form of representing scanned artefacts is as a textured triangle mesh. A serious 3D infrastructure, however, requires a more diverse set of shape representations. Scanned cuneiform tablets, for instance, usually have a multitude of view-dependent textures or under different lighting conditions, whereas computer tomography produces a volumetric "image" of what is inside, e.g., an Egyptian mummy [BSly]. This requires that new shape representations can be integrated with the viewer, e.g., loaders for new file formats, and new 3D rendering modules, for example for the BTF models from Bonn (see Fig. 12).

Cuneiform viewer module: The cuneiform tablets (see Fig. 6,2.5) so far require a viewer that is inherently 2D (shown in Fig. 4), and not 3D. It would be highly desirable, however, to overcome this artificial separation: The 2D relighting application should actually be embedded in 3D (as another low-level module), because only then the 2D view can be shown together with the 3D model. The way to go would be to generalize the 2D application in the sense that the 2D content is renderd on a rectangle floating in 3D that is part of the normal scene graph. – The only remaining issue is that advanced 2D viewers today also heavily use the GPU. A solution might be upcoming GPU frameworks like CUDA from Nvidia, which have mechanisms for management of the GPU resource.

Template extensible: Whereas the main use of the 3D kiosk is to let a visitor explore one single artefact, there is a wide range of possible presentation scenarios. Users might pick one among many artefacts from a shelf, or from a digital replica of the museum room, or even from a historic scene digitally rebuilt in order to contextualiize the cultural artefacts by showing them in their historic sur-

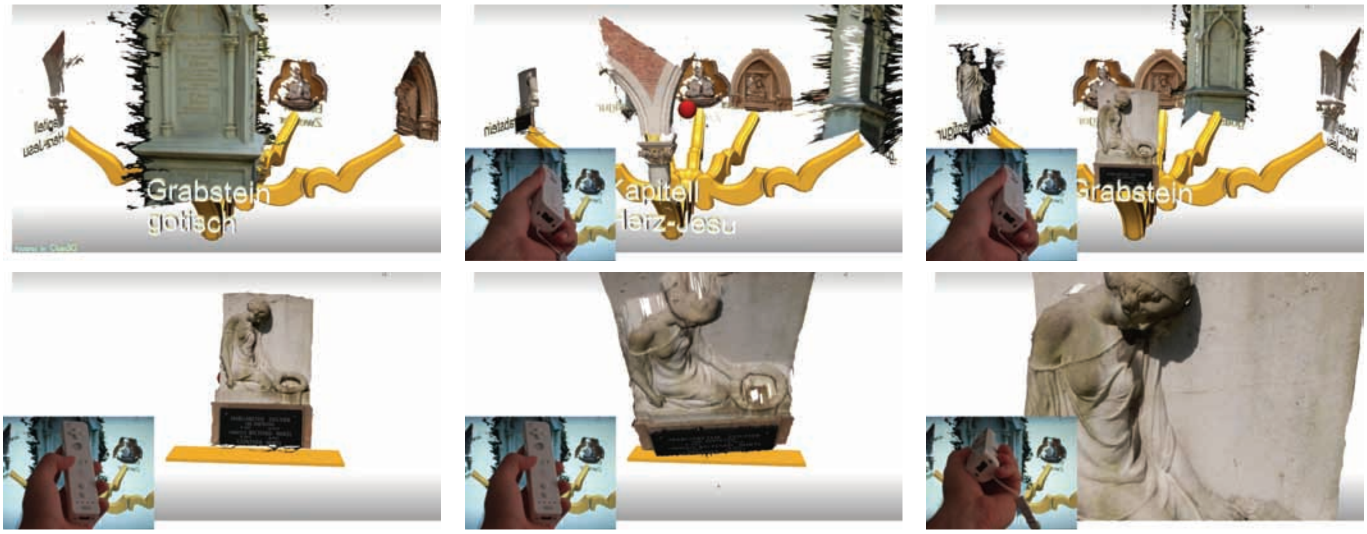


Figure 15: *Example of a 3D presentation. Upper row: The user can rotate the chandelier horizontally. Lower row: One object is chosen for inspection, and the user can orbit around it. – The objects shown are not high-quality artefacts but only examples.*

roundings. This flexibility shall become possible through *customizable presentation templates*, very basic 3D scenes with objects that have a reasonable pre-defined behaviour and whose appearance (geometry+texture) can be configured via drag-and-drop. Curators can download these presentation templates from suitable websites.

3D modeling of ad-hoc geometry: Sometimes ad-hoc objects are needed for a presentation. Static objects could be created photogrammetrically from digital photographs using the Epoch Webservice [VG06]. However, this is not applicable in all cases. To let users generate simple objects, e.g., extruded 2D contours, a very simple 3D modeling tool should be part of the authoring software. This tool is comparable to the vector-based diagram editor included in, e.g., Microsoft Powerpoint. And just like Powerpoint it should allow to animate these diagrams by animating the object parameters. This way a 3D stone wall could vary its position and (x,y,z)-extents very much like a line segment in a 2D drawing can.

Non-monolithic: From a software point of view the presentation viewer shall behave like a component, rather than like a stand-alone application. The reason is re-usability: Along with the 3D presentation additional textual information might have to be displayed, a HTML page or a pdf document. It shall even be possible to integrate the 3D presentation with another application that has its own GUI, such as a numeric simulation or a database front-end. The consequence is that the 3D presentation viewer shall requires not much more than a 3D window to render into; another consequence is that it does not provide a sophisiticated 2D GUI with a menu hierarchy (like MS Powerpoint has). It should be possible, though, to later add a 2D GUI with a menu.

Developer Levels: We envisage a hierarchy of users of our systems. Each level requires more knowledge and, thus, will reach out to a smaller community:

Level 0: End-user who consume the 3D-presentations
Level 1: Authoring of presentations: 3D-GUI, drag&drop
Level 2: Authoring of presentation templates: Scripting
Level 3: Extension developers: C++ programming

Users on levels 1-3 are creative people, the DCC providers, which stands for *digital content creation*.

6. The Epoch Viewer

Our solution is to based on the combination of the OpenSG scene graph system with the GML scripting language [Hav05].

We have developed a series of GML scripts for 3D modeling, for presentation templates, and for particular presentations. The first phase of our work concentrated on providing OpenSG with the functionality needed, and on making it accessible via GML in a suitable fashion. Next we have begun to create a number of example presentations on this basis. The next step, which will start soon, is to revise and refactor the GML code for those presentations. The goal is to distill a set of basic GML components out of these presentations, in order to produce a GML framework that will be useful for all sorts of presentations. The final step will be to create a conventional GUI and menu system, which makes use of the 3D presentation as a component.

6.1. Input device: The tangible proxy object

By far the most intuitive 3D input device is a 3D object. The idea of the *proxy object* is that the virtual object displayed on the screen moves exactly like the real object that the user holds in his hands. Ideally, the user can move and rotate the object, and the digital artefact is in exact sync. Note, however, that we want to map the 6-DOF pose directly, not in a mirror fashion, so that when the user stretches out the hand with the proxy object the virtual object also goes farther away. It does not come closer as would be the case with a mirror. – Technically, the problem is to determine the 6-DOF pose of the proxy object. We have experimented with the three technologies shown in Fig. 14.

First technology: ARToolkit. We have tried camera-based tracking using the ARToolkit from www.artoolkit.org. With one marker

Figure 16: *Authoring a presentation with BTF-models. They are loaded and rendered using an OpenSG extension module that makes use of advanced vertex and pixel shaders. Models are courtesy Gero Müller and Reinhard Klein, Univ. Bonn, Germany*

per side of a cube and a single camera we had serious robustness problems: Whenever a finger only *touched* the black boundary of the marker the recognition algorithm of ARToolkit broke. Consequently we have made the cube a bit larger and used an array of 2×2 markers. This seems to be a good compromise between robustness and processing load, as the latter affects the recognition speed and, thus, the frame rate. It is quite unlikely that the user occludes all four markers of one side at the same time, and usually more than one side is visible.

Second technology: Accelerometer. ARToolkit markers have a very technical appearance which we wanted to avoid. The position tracking also created slight problems since users tended to move the 3D object out of the frustum: To inspect the object they took it very closely until the view showed only a detail of the surface; but then, in order to rotate it, they took it from one hand to the other, thereby actually moving it quite a bit to the right or the left. This made the object suddenly disappear, and the users got lost and felt uncomfortable struggling to bring the object back.

Consequently we tried to get along by using only the orientation (acceleration) information. With a object held still, gravity causes the strongest acceleration, so the downwards direction can be robustly detected. Roll information around the gravity vector, though, can not be detected, so it can not be decided whether the user points north, east, or west. So we used the accelerometer information only for relative motion (spinning speed). This worked well and robustly.

The accelerometer is a standard electronic device and quite cheap (15 Euros). It can easily be connected to a serial or USB port. Due to its small size it can also be put inside another object, e.g., one that resembles a cultural artefact. This looks much better in a museum than ARToolkit markers.

Third technology: Nintendo Wiimote. The controller of the Nintendo Wii, the *Wiimote*, communicates with standard Bluetooth. Free software tools exist to decode its protocol, e.g., Kenner's GlovePIE [Ken]. The Wii can deliver also position information, as it contains an optical sensor that, when pointing towards a certain configuration of LEDs, determines the pose relative to the LEDs. The Wiimote is a mature device and quite robust to use, which made it our preferred test device, despite its non-museal appearance.

6.2. The 3D Presentation

A first example of a 3D presentation is shown in Fig. 15. The user sees a nobject selection menu that is shaped like a chandelier. With a slight rotation of the Wiimote to the left or the right the chandelier begins as well to rotate smoothly, showing the next object in the respective direction. By tilting the Wiimote upwards the close-up view is activated: The chandelier gradually moves away and the chosen object comes close until it fills the view.

One of our goals was that users can always keep track of what is going on. There are no abrupt transitions and we have taken care that all motions are smoothly animated. Before the chandelier moves away, it retracts; when the close-up inspection is finished, the chandelier appears again and unfolds, see Fig. 18. The close-up object can be inspected in detail: With the two DOFs of the Wiimote (rotate L/R and U/D, for left/right, up/down) it is only possible to orbit around the object center in a fixed distance: In terms of Euler angles, L/R determines the azimuth and U/D the elevation of the object.

Figure 18: *The "chandelier" smoothly retracts before moving away, as a clear signal that close-up inspection begins*

We have experimented also with a combined interaction mode: The elevation must be clamped, e.g., to $[-70, 70]$ degrees to avoid the gimbal lock. When the elevation is maximal or minimal, a further increase or decrease makes the object come closer or get farther away, respectively. – Users usually reported a good feeling of control in this combined mode. The problem was only that they found it uncomfortable: It is apparently more convenient, or natural, to first navigate (orbit) to a particular spot on the surface, and then to adjust the distance to this spot. Our mode required to first adjust the distance, and then orbit over the surface.

Note that the objects in Fig. 15 are low quality reconstructions, generated photogrammetrically from a single range map and decimated to 10K vertices using Michael Garland's quadric-based simplification tool `qslim`. A much higher rendering quality can be obtained using the BTF-rendering module for OpenSG from Gero Müller and Reinhard Klein (Univ. Bonn). A BTF provides much more surface detail as it approximates the BRDF with much more than only a single texture value per surface point. Especially small and shiny BTF surface parts are brought out by the headlight much more clearly. Fig. 16 can only deliver part of the experience to hold a shimmering object virtually in his own hands.

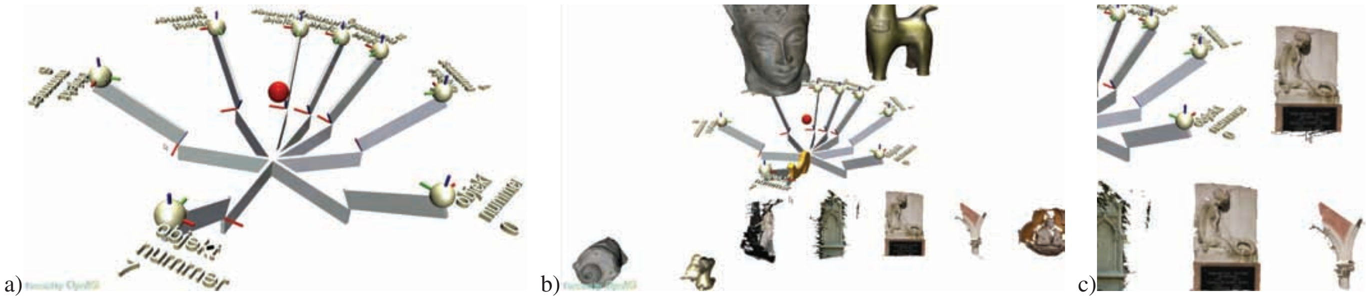


Figure 17: *Authoring via drag-and-drop. a) A blank layout consists of many drop targets. b) The "model bar" is filled via drag-and-drop from the file system with 3D models, 2D images, and 1D text strings. c) Objects are dragged interactively from the model bar to drop targets in the 3D scene, where they automatically align to the local coordinate frame.*

6.3. Authoring of a 3D Presentation

Figure 19: *Presentation layout: Parametric shelf. All the boxes can be dragged to adjust the spacing, and pushing the red spheres inserts a new row or column of planks.*

Figure 20: *Presentation layout: Chandelier. Even when densely populated, the selected object sticks out clearly due to the uneven circular spacing.*

We have experimented with several sorts of layouts. The first idea was a rack or shelf (Fig. 19) where the spacing can be interactively adjusted to match the sizes of the artefacts. Fig. 20 shows our chandelier-like design. Its special feature is that it rotates non-linearly in order to clearly highlight the object that can be chosen for the detailed inspection.

6.4. Authoring via Drag & Drop

The main idea is that our layouts are almost completely composed of so-called *drop targets*. Fig. 17a shows such a "blank" layout. All the boxes and spheres are successively replaced. Three different types of objects are supported: **3D models** (Collada .dae, Wavefront .obj, Stanford .ply, OpenSG .osb, etc), **images** (.png or .jpg, immediately applied to texture a quadrangle), and **character strings**, which are rendered as true 3D text. The replacement proceeds in two steps:

- **Filling the model bar:** The user drags an object from the file system (the Windows Explorer) to the 3D window where it appears in a row along the lower border of the 3D window, the *model bar* (Fig. 17b)
- **Replacing the drop targets:** Individual models can be dragged interactively from the model bar to drop targets in the 3D scene (Fig. 17c). Whenever dragging the object over a suitable drop target the object temporarily snaps and aligns with this target. The user can decide whether to leave it there (mouse release), or continue to drag it elsewhere.

Note that objects can not be dragged immediately from the file system to a drop target in the scene; the model bar always acts as an intermediate storage. The reason is that with larger models there is a noticeable delay until the object is loaded and appears in 3D under the mouse pointer. Users would instinctively think that the dragging operation has failed, and stop dragging – only to see that after a short while the object appears *somewhere* in the scene. This was perceived so frustrating that we decided to introduce the model bar.

Another thing that has proven successful was that when loading an object we immediately show a temporary geometry, a sphere, that is replaced by the true object as soon as its loading is finished. We do not, however, use the sphere for immediate drag-and-drop because of size issues: The temporary sphere can not reflect the true size of the object that is being loaded, simply because the bounding box of this object is only available after it has been loaded.

We plan to solve this issue using the so-called *Collada light* 3D format: The (lightweight) Collada XML file contains only semantic and metadata information, in particular the bounding box, and it references another (heavy) binary file that contains the 3D data, for instance a huge U3D file with compressed triangle data.

6.5. Modeling Included: Creating Ad-Hoc Geometry

The design of the arms of the chandelier was based on a small sketch on a piece of paper that, in the museum scenario, would have come from a curator or from a graphics designer (see Fig. 21). The sketch was photographed, the photograph was perspectively corrected by marking four points and then it was loaded into the GML based modeller. The modelling proceeds in a very simple rubber-band fashion using a control polygon (Fig. 23): Clicking on the polygon boundary inserts a new ball. Balls can be freely dragged around on the construction plane. Just clicking on a ball toggles its red/green status: green balls control a B-Spline, red balls are corners. Clicking

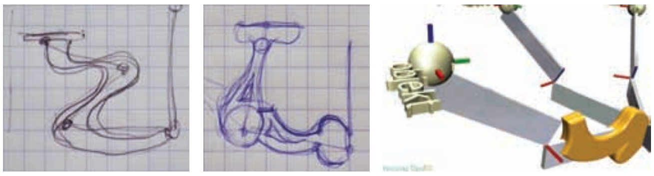


Figure 21: *Two designs for the arm of the chandelier. The ad-hoc geometry created from the sketches is being dragged onto the drop target. As soon as the mouse button is released it is copied on all instances of the arm.*

on the blue ball extrudes the polygon. The extrusion profile can also be adjusted, but this is not shown here. – The ad-hoc geometry thus created is then dragged into the authoring toolkit.

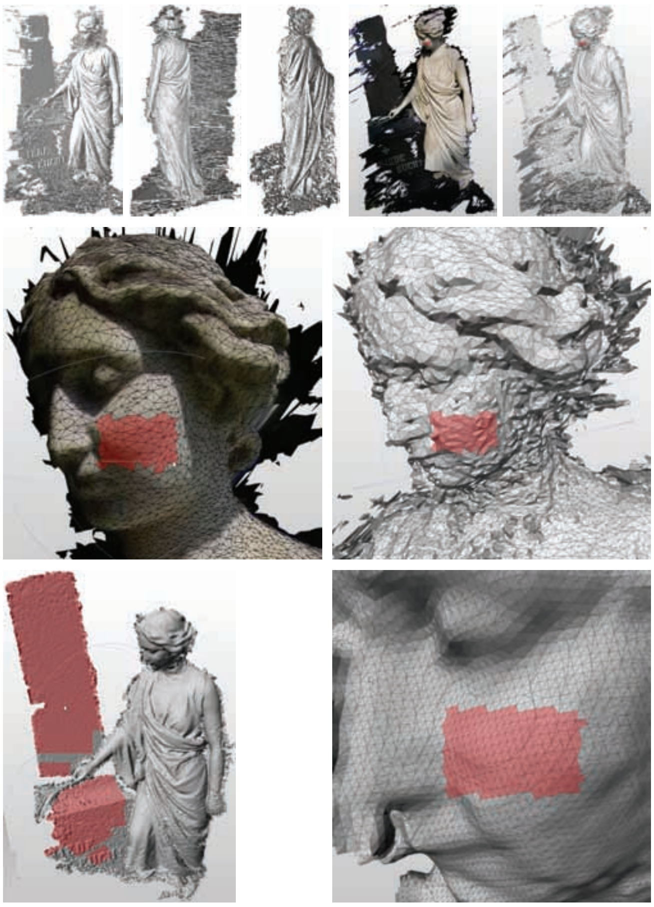


Figure 22: *The shape markup problem. The* ***cheeks*** *of the statue are a concept that exists independently from the particular shape representation that is used. The problem is how to store the shape markup in a sustainable way such that it is robust against remeshing and mesh processing of the various input data (first row) that are used to obtain the final model (bottom right).*

7. Conclusion and Future Work

The Epoch viewer presented in this paper leaves room for further improvement; in fact, it is still work in progress, since by far not all features from the feature wish list could be realized. But although our results are far from perfect we claim that the foundation is sound with respect to the feature wish list from section 5, with which we would also highlight some areas for further research:

- **Drag-and-Drop Authoring:** The drop targets work extremely well in the authoring mode. Future layouts will have also drop targets for images (backdrop) and text. The model bar should be part of a 2D GUI, though.
- **Versatile Hardware Setups:** Any input device that uses the channel concept can connect via socket. However, this is not a solution for input devices that require very much bandwidth such as, e.g., cameras.
- **Easy 3D navigation:** The tradeoff between comfort and control must be improved further by employing more intelligent camera behaviours. 3D navigation in the authoring mode must be improved as well.
- **CH Integration:** As soon as Collada files offer more standard (meta-)information, this information should be available for 3D presentations. Examples: The textual caption of the 3D model, and information links embedded in the surface.
- **CH Sustainability:** Clearer separation between framework and individual layout, so that the layout style of a presentation can be exchanged like in MS Powerpoint.
- **Low-Level extensible:** Accomplished in principle, but unclear for visualisations that are genuinely 2D.
- **Cuneiform viewer module:** This is genuinely 2D and, thus, still open.
- **Template extensible:** Accomplished in principle but now needs many more examples
- **3D modeling of ad-hoc geometry:** It is unclear how sophisticated the 3D modeling should be: Should it be like Google Sketchup or significantly less powerful? – Very interesting though would be the use of animated parametrized models as 3D-illustrations.
- **Non-monolithic:** This is definitely accomplished. The presentation viewer can in fact be embedded as a component into any application providing an OpenGL window. All program functions can be accessed through GML scripts that can even be synthesized at runtime.
- **Developer Levels:** A proper documentation of the scripting facility is the only thing that is required for others to develop interesting presentation templates.

The remaining problems can be seen as symptoms of underlying much more fundamental issues (that are alluded to in Fig. 22). The articles on the *seven open research questions* [HF07, BFH*07] for example identify some areas where radically new solutions and approaches need to be developed before we are able to manage really large numbers of valuable Cultural Heritage assets in digital libraries reliably and sustainably.

Acknowledgements

We would like to thank prof. Karel Van Lerberghe and Hendrik Hameeuw for the given insight into the study of cuneiform tablets and the extended dissemination of the acquisition system.

References

[Ben07] BENDELS G. H.: *Methods for 3D Geometry Processing in the Cultural Heritage Domain*. PhD thesis, University of Bonn, Germany, 2007.

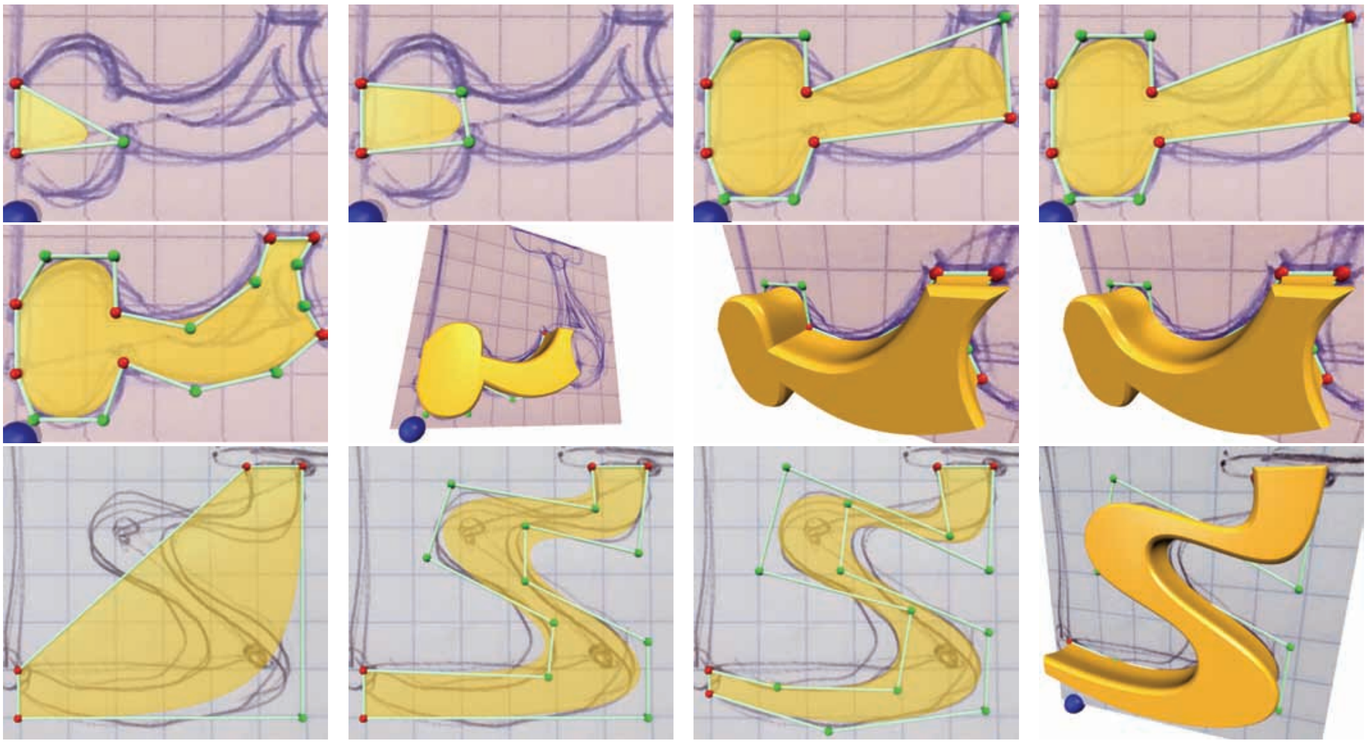


Figure 23: *Creating ad-hoc geometry from a sketch. Row 1: A rubber band triangle is successively expanded. Row 2: The profile is extruded, polygon vertices can be sharp (B-spline) or smooth (corner). Row 3: Most users understand quickly how to place vertices in a way such that the resulting raw shape can be adjusted efficiently.*

[BFH*07] BUSTOS B., FELLNER D., HAVEMANN S., KEIM D., SAUPE D., SCHRECK T.: Foundations of 3d digital libraries: Current approaches and urgent research challenges. In *Proceedings of the First International Workshop on Digital Libraries Foundations (DLF1), Vancouver, British Columbia, Canada, June 23, 2007* (2007), DELOS Network of Excellence on Digital Libraries.

[BSly] BBC, SILICON GRAPHICS (SGI): Mummy: Inside story, 2004 July. Exhibition at the British Museum London, http://www.thebritishmuseum.ac.uk.

[CGS06] CHANG T., GOESELE M., SEIDEL H.-P.: Mesostructure from specularity. In *CVPR* (2006), pp. 1825–1832.

[CMFP02] COSTAGLIOLA G., MARTINO S. D., FERRUCCI F., PITTARELLO F.: An approach for authoring 3d cultural heritage exhibitions on the web. In *SEKE '02: Proc. 14th int'l conference on Software and knowledge engineering* (New York, NY, USA, 2002), ACM Press, pp. 601–608.

[CPCS06] CALLIERI M., PONCHIO F., CIGNONI P., SCOPIGNO R.: Easy access to huge 3d models of works of art. In *Fourth Eurographics Italian Chapter 2006* (Graz, Austria, 2006), Fellner D., (Ed.), Graz University of Technology, Eurographics Association, pp. 29–36. Catania, 22-24 February 2006.

[CPCS08] CALLIERI M., PONCHIO F., CIGNONI P., SCOPIGNO R.: Virtual inspector: a flexible visualizer for dense 3d scanned models. *IEEE Computer Graphics and Applications 28*, 1 (Jan.-Febr. 2008), ??–??

[DMK03] DEGENER P., MESETH J., KLEIN R.: An adaptable surface parametrization method. In *The 12th International Meshing Roundtable 2003* (September 2003).

[DNvGK97] DANA K., NAYAR S., VAN GINNEKEN B., KOENDERINK J.: Reflectance and texture of real-world surfaces. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings* (1997), 151–157.

[epi07] EPISODE E. F.: European digital library to contain all knowledge. (Belgian television), 2007. www.euronews.net/index.php?page=futuris&article=444863.

[Hav05] HAVEMANN S.: *Generative Mesh Modeling*. PhD thesis, Institute of Computer Graphics, Faculty of Computer Science, Technical University Braunschweig, Germany, November 2005.

[HBD07] HAHN D. V., BALDWIN K. C., DUNCAN D. D.: Non-laser-based scanner for three-dimensional digitization of historical artifacts. *Applied Optics 46*, 15 (May 2007), 2838–2850.

[HF04] HAVEMANN S., FELLNER D. W.: Generative parametric design of gothic window tracery. In *Proc. VAST 2004: The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage* (Brussels, Belgium, 2004), Cain K., Chrysanthou Y., Niccolucci F., Silberman N., (Eds.), Eurographics, pp. 193–201.

[HF07] HAVEMANN S., FELLNER D. W.: Seven research challenges of generalized 3d documents. *IEEE Computer Graphics and Applications 27*, 3 (2007), 70–76. (special issue on 3D documents).

[HS03] HERNÁNDEZ C., SCHMITT F.: Silhouette and stereo fusion for 3d object modeling. In *3DIM 2003, 4th International*

Conference on 3D Digital Imaging and Modeling (2003), pp. 46–53.

[HSKF06] HAVEMANN S., SETTGAST V., KROTTMAIER H., FELLNER D.: On the integration of 3d models into digital cultural heritage libraries. In *VAST 2006 Project Papers* (Nicosia, Cyprus, Nov 2006), Ioannides, Arnold, Niccolucci, Mania, (Eds.), Epoch public., pp. 161–169.

[Ken] KENNER C.: Glove programmable input emulator (glovepie). carl.kenner.googlepages.com/glovepie.

[Lau94] LAURENTINI A.: The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Mach. Intell. 16*, 2 (1994), 150–162.

[MBK05] MÜLLER G., BENDELS G. H., KLEIN R.: Rapid synchronous acquisition of geometry and btf for cultural heritage artefacts. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)* (November 2005), Eurographics Association, Eurographics Association, pp. 13–20.

[MBR*00] MATUSIK W., BUEHLER C., RASKAR R., GORTLER S. J., MCMILLAN L.: Image-based visual hulls. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 369–374.

[MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2001), ACM, pp. 519–528.

[MMK04] MÜLLER G., MESETH J., KLEIN R.: Fast environmental lighting for local-pca encoded btfs. In *Computer Graphics International 2004 (CGI 2004)* (June 2004), IEEE Computer Society, pp. 198–205.

[MSK06] MÜLLER G., SARLETTE R., KLEIN R.: Data-driven local coordinate systems for image-based rendering. *Computer Graphics Forum 25*, 3 (September 2006).

[MVSL05] MUDGE M., VOUTAZ J.-P., SCHROER C., LUM M.: Reflection Transformation Imaging and Virtual Representations of Coins from the Hospice of the Grand St. Bernard . In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST* (Pisa, Italy, 2005), Mudge M., Ryan N., Scopigno R., (Eds.), Eurographics Association, pp. 29–39.

[MVW*06] MÜLLER P., VEREENOOGHE T., WONKA P., PAAP I., GOOL L. V.: Procedural 3D Reconstruction of Puuc Buildings in Xkipché. In *International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage* (Nicosia, Cyprus, 2006), Ioannides M., Arnold D., Niccolucci F., Mania K., (Eds.), Eurographics Association, pp. 139–146.

[MWGA06] MALZBENDER T., WILBURN B., GELB D., AMBRISCO B.: Surface Enhancement Using Real-time Photometric Stereo and Reflectance Transformation. In *Symposium on Rendering* (Nicosia, Cyprus, 2006), Eurographics Association, pp. 245–250.

[VG06] VERGAUWEN M., GOOL L. V.: Web-based 3d reconstruction service. *Machine Vision Applications*, 17 (2006), 411–426.

[VWV06] VERBIEST F., WILLEMS G., VAN GOOL L.: Image-based rendering for photo-realistic visualization. *Virtual and Physical Prototyping (Taylor & Francis) 1* (2006), 19–30.

[WVM*05] WILLEMS G., VERBIEST F., MOREAU W., HAMEEUW H., VAN LERBERGHE K., VAN GOOL L.: Easy and cost-effective cuneiform digitizing. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST),* (Pisa, Italy, 2005), pp. 73–80.

EPOCH Conference on Open Digital Cultural Heritage Systems (2008)
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

Tools for Populating Cultural Heritage Environments with Interactive Virtual Humans

D. Arnold[1], J. Glauert[2], V. Jennings[2], B. Kevelham[3], N. Magnenat-Thalmann[3], Jonathan Maïm[4], D. Maupu[4], G. Papagiannakis[3], K. Rodriguez-Echavarria[1], D. Thalmann[4] and Barbara Yersin[4],

[1]University of Brighton - Brighton, United Kingdom
[2]University of East Anglia - Norwich, United Kingdom
[3]UNIGE - Geneva, Switzerland
[4]EPFL - Lausanne, Switzerland

Abstract

Modern 3D VR systems rely heavily on the interplay of heterogeneous technologies. Because of this inherently interdisciplinary character, VR domain can be viewed as a melting pot of various technologies which although complementary are non-trivial to put together. Frameworks can be used to address this challenge as they offer advantages such as reusability of components, as well as easiness of replacements, extensions, and adaptations. Hence, this paper presents developments within the EPOCH project, in particular the Characterize NEWTON, to improve and release frameworks that support the incorporation of avatars in interactive real-time 3D VR systems. The purpose is to enable avatars to be interactive and to react to model metadata; thus adding realism and engaging the user's interest. This vertical middleware offers the advantage to be based on open source generic frameworks, such as OpenScenegraph and OpenSG as well as offering complementary functionalities.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Three-Dimensional Graphics and Realism]: Virtual reality I.3.4 [Graphics Utilities]: Graphics packages

1. Introduction

Once Cultural Heritage (CH) data has been captured, stored, and interpreted; CH professionals usually turn their attention to the display of their findings in interactive 3D Virtual environments. For this, it is not only necessary to create a 3D scene but also to add realism to it by including virtual humans, often referred to as avatars.

Modern 3D VR systems rely heavily on the interplay of heterogeneous technologies. Because of that inherently interdisciplinary character, VR domain can be viewed as a melting pot of various technologies which although complementary are non-trivial to put together. Many of those technologies gain a lot of individual R&D interest but still it is not generally understood and there are barely any accepted guidelines and approaches in the matter of integration of those functional artifacts under the roof of one consistent framework. In other words, there are many masterpieces of atomic technologies but still there is a lack of a well understood and generally accepted strategy for putting them up so they would constitute the whole bigger than the simple sum of its parts. The missing element is an open source framework which would curb the complexity and make the resulting system machinery a consistent and seamless unity, leaving at the same time open handles and hooks for replacements and extensions.

It becomes visible that the complexity of such systems reaches the levels that cannot be handled anymore efficiently by methodologies and technologies of today. Object oriented toolkits, knowledge of well established architectural and development patters plus human skills and experience still help and do the job but in order to stay on the cutting edge of tomorrow development speed, massive reusability of components, easiness of replacements, extensions, adaptations, reconfigurations and maintenance must be addressed. These are exactly the features offered by the frameworks.

Within the context of the EPOCH project, in particular the Characterize NEWTON, we aimed *to improve and release frameworks that support the development of interactive audio-visual real-time 3D VR systems*. In particular, environments which feature real-time virtual character simulation with state-of-the-art clothing. For this, the main challenges undertaken by this research were: elaboration of methodology, guidelines, architectural, design and behavioral patterns leading to the construction of vertical middleware frameworks. This resulted in the release of an easy to use set of tools to incorporate avatars that react to model metadata adding realism and interest for users. This vertical middleware offers the advantage to be based on open source generic 3D graphic frameworks, such as OpenScenegraph [Ope08a] and OpenSG [Ope08b] as well as offering complementary functionalities.

The following sections will describe the frameworks resulting from this work: a) the release as open source of the VHD++ kernel and plug-in and b) the improvement of the UEA Scene Assembly

Toolkit. The latter supports the use of natural language interaction. Furthermore, two applications will be presented illustrating the environments and efficient virtual human simulation which could be created by using these tools. A brief overview on the usability and acceptability for the type of application produced by this framework is also described.

2. VHDPlus

UNIGE and EPFL have been actively involved in the EPOCH project contributing with virtual human simulation technologies as part of the EPOCH Common Infrastructure in both showcases as well as the Characterize NEWTON projects by adapting and releasing their core platform vhdPLUS as open-source tool for the cultural heritage community. The vhdPLUS Development Framework is a modern, fully component oriented simulation engine and software middleware solution created by and reflecting many years of the R&D experience of both the MIRALab, University of Geneva and VRlab, EPFL labs in the domain of VR/AR and virtual character simulation [PPM*03].

vhdPLUS is a highly flexible and extensible real-time framework supporting component based development of interactive audio-visual simulation applications in the domain of VR/AR with particular focus on virtual character simulation technologies (see figure 1). It relies heavily on multiple, well established Object Oriented design patterns, uses C++ as the implementation language and Python as a scripting language.

Figure 1: *Screenshots from the final 3D interactive real time virtual heritage simulations based on vhdPLUS*

vhdPLUS has been released as open source as part of the EPOCH-NEWTON Characterize activities. In support of this release, and as per the Common Infrastructure activities, we have created several resources that should lead to a more effective use of vhdPLUS for users in general, but especially for those involved in the Characterize activities. Some of the components in the vhdPLUS version include:

- OpenSceneGraph rendering.
- OpenSG based rendering for static objects.
- VRML97/HANIM1.1 low level parser library.
- Helper library for Virtual Human control: libvhdOSGExt
- Configuration of vhdPLUS through XML files

To allow the interaction between vhdPLUS and OpenSG a new Service (a vhdPLUS plug-in) has been written that allows for the rendering and placement of geometry through OpenSG. Furthermore a library has been provided for the loading and animation of HANIM1.1 virtual humans and we included a service template with basic code to demonstrate the inclusion of new services into vhdPLUS.

A further explanation of the use of XML in combination with vhdPLUS has been given on the website: [vhd08b]. We have also included basic doxygen generated documentation showing the structure of the various components that make up vhdPLUS as well as a paper demonstrating the use of VHD++ (and therefore vhdPLUS) in cultural heritage contexts [NMTCY07] and [MTP06]. Based on various user inputs, UNIGE has updated the basic building environment and scripting tools for easier adoption of the framework. EPFL has proceeded with the addition of specialised to virtual human simulation math library.

Finally, following a recommendation from the last EPOCH review, clear references and acknowledgments to EPOCH have been added on the dedicated web-site. Since 2007, more than 200 downloads have illustrated the interest of the open source community on the above framework.

2.1. Availability

vhdPLUS has been made available through Sourceforge [vhd08a], and has been released under the LGPL2 license. It is accompanied by a website [vhd08b] in which we have made available a number of documents [PPM*03] and [MTP06] detailing the full structure of vhdPLUS as well as its capabilities and uses. Especially [Pon04] gives a full description of VHD++ (vhdPLUS's parent framework) including information plug-in structure (called Services in vhdPLUS) as well as XML initialisation.

3. UEA Scene Assembly Toolkit

The University of East Anglia (UEA) has been, as part of the Characterize NEWTON, involved in improving a toolkit for 3D scene creation. This toolkit supports the user through an assembly pipeline by offering the following components:

1. Terrain Converter (see figure 2): converts terrain data from a regular grid to a ground model.
2. Avatar Research Platform (ARP) toolkit: supports the design and animation of Virtual Humans capable of speech and sign language.
3. Scene Assembler: used to combine objects in the scene using scriptable operations and manual tweaking. For this, every command is stored in an XML script format for rapid semiautomatic assembly of scenes.

The Avatar Research Platform (ARP) toolkit includes tools for conventional mesh attachment and design. Some of its features include:

- Advanced tools for facial morph creation (as shown in figure 3).
- Bespoke avatar creation for procedural animation (UEA Animgen).
- Imports industry standard Maya and FBX.
- Exports to Maya and FBX.

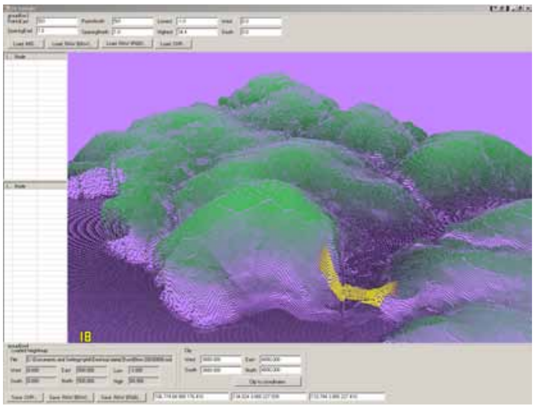


Figure 2: *Scene Assembly Toolkit: Terrain Converter component*

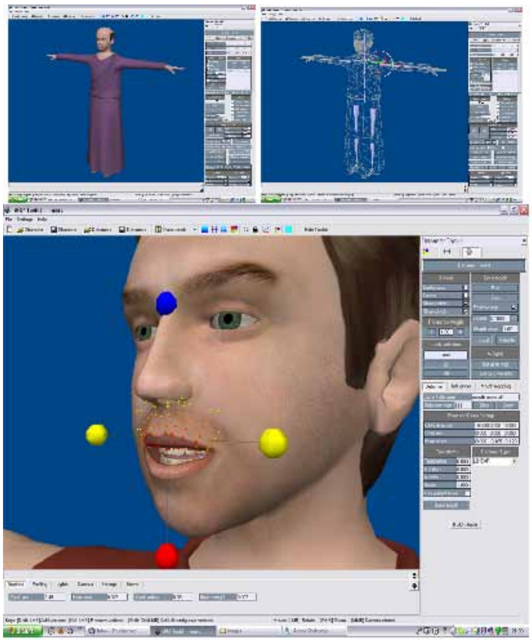


Figure 3: *Screenshots of UEA Avatar Research Platform (ARP)*

- ARP format used by EPOCH Scene Renderer.

A propietrary extension for COLLADA [Col08] standard is used as the format for exporting from the ARP toolkit. The Scene Assembler can import virtual humans in this format into scenes. To work smoothly, it has been required to resolve some issues in OpenSG with the dynamic nature of virtual humans, such as memory overheads, rendering speeds, synchronization of audio and visual components important for speech. This was done by utilising "deprecated" interfaces that allow low-level use of OpenGL. Furthermore, interfaces have been provided for attaching hyperlinks to objects so that users can access related resources, for example websites.

As illustrated in figure 4, the Scene Assembler tool creates a final scene by adding simple static geometry objects from popular modelling software (3DSMax, Maya, etc.). In addition, billboard objects defined by text are used to represent trees, lamp posts, and other street furniture. Objects from a group can be "planted" with a single click, especially useful for rapidly adding vegetation. Then, individual virtual humans are placed in the scene with animations as shown in figure 5. The Scene Assembler tool can export scenes in .OBJ format (static objects only) or COLLADA with extensions. Rendering of the final scene is based on the open source OpenSG rendering framework and utilizes occluder culling techniques for rapid animation.

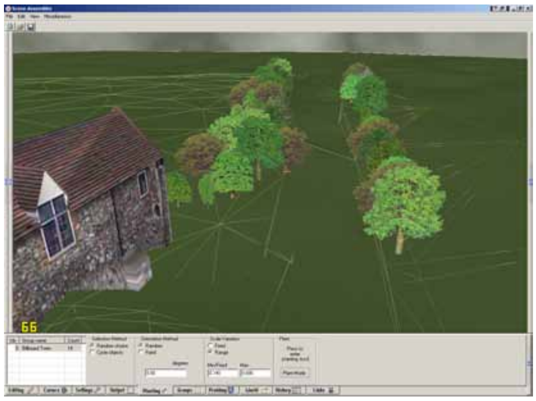


Figure 4: *Scene Assembly Toolkit: Scene Assembler component*

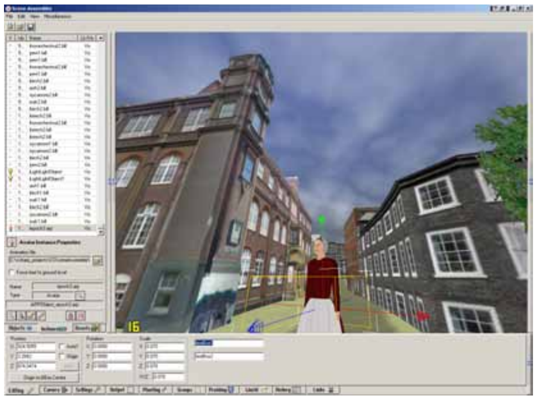


Figure 5: *Inclusion of virtual avatars in a 3D scene*

Within this framework, the Natural Language Processing and Generation module developed by the University of Brighton provides natural language interactivity within the environment. This functions as a question/answer system using predominantly natural language. For this, language technology automates the structuring and querying of heterogeneous and semi-structured information in databases which are structured within the CIDOC-CRM ontology [CID08]. The CIDOC Conceptual Reference Model (CRM) provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation.

In addition, the system uses semantic similarity for producing appropriate answers to user's queries. This similarity is computed using dynamic programming where deletion, insertion and substitution are given a cost. WordNet provides a measure of the cost for substituting one word for another. As a result, the user query Q is compared with a set of n predefined queries. If the most semantically related Q-Qi pair is above a given threshold of confidence, the corresponding answer Ai is selected. Otherwise an Eliza-type of interaction is adopted.

Finally, the potential for integration of the resources and 3D environments produced by both of these frameworks is achieved by using the common open source generic frameworks OpenSG. For this, vhdPLUS includes an OpenSG rendering plug-in and UEA framework is based on the OpenSG rendering framework.

4. Interactive 3D Virtual Environments

In order to illustrate the environments and efficient virtual human simulation produced by both frameworks, two different real-time interactive 3D VR systems environments will be presented in the following sections. Both of these applications recreate a city or town populated by virtual humans for visitors to explore and gain information on the history of the place in an engaging and entertaining way.

4.1. Reviving the ancient city of Pompeii

Its main goal is to simulate in real time a crowd of virtual Romans exhibiting realistic behaviors in a reconstructed district of ancient Pompeii as illustrated in figure 6.

In an offline process, the city is first automatically reconstructed and exported into two different representations: a high-resolution model for rendering purpose, and a low-resolution model labeled with semantic data. Second, the city is populated with crowds of convincing virtual Romans presenting several different behaviors, and thus offering a realistic and varied simulation.

The city reconstruction is based on a CGA Shape grammar with nine different facade designs derived from archaeological drawings. It contains 4 levels-of-detail, 16 different wall types and three roof styles. Of this grammar 16 variations were automatically generated by combining the facades and roofs with specifically designed color palettes. This number could be arbitrarily increased, but practical aspects of the rendering limited the usable number of materials. The CGA Shape grammar has proven its great flexibility, for instance during the optimization of the levels-of-detail for the rendering process.

There are several buildings in the city model where virtual Romans can enter freely. Some of them are labeled as shops and bakeries, and the characters entering them acquire related accessories, e.g., oil amphoras or bread. These accessories are directly attached to a joint of the virtual character's skeleton, and follow its movements when deformed. We can attach accessories to various joints, depending on their nature. In Pompeii, this variety is illustrated with the amphoras: rich people leave shops with an amphora in their hand, while slaves leave shops carrying them on their heads.

The idea of rich and poor districts is based on age maps that were provided by archaeologists taking part in the EPOCH project. These maps show the age of buildings in the city. Although we do not yet have the building textures to visually express this kind of differences, we have decided to install the rich Roman templates in the most recent districts, while poor people have been established in old buildings. From this, virtual characters know where they belong and while most parts of the city are accessible for everybody, some districts are restricted to a certain class of people: rich Romans in young areas and slaves in poor zones.

As for the crowd, seven human templates have been exploited: a couple of nobles (one male and one female), a couple of plebeians, another couple of patricians and finally, a legionary. These seven templates are instantiated several hundred times to generate large crowds. To ensure a varied and appealing result, per body part color variety techniques are exploited. The resulting application is presented in [MHY*07].

Figure 6: *Crowds of virtual Romans in a street of Ancient Pompeii*

4.2. Touring the town of Wolfenbuttel

This application uses the UEA Scene Assembly Toolkit. It recreates Wolfenbüttel as it once stood during the seventeen century. This town sits on the Oker river in Lower Saxony-Germany, just a few kilometres south of Braunschweig. Within the 3D interactive application, the user navigates the scene accessing information about important buildings. For this, important landscapes in the scene are annotated with meta data so that visitors can explore related websites or other resources during a virtual tour.

To build the application, 3D modelling packages were used to model the most important buildings of the town. These models and other generic models of houses were imported to the Scene Assembler and the final scene was exported in the COLLADA [Col08] format. The showcase demonstrator was used for real time rendering of the imported COLLADA file using OpenSG [Ope08b], and provides user interfaces for navigation through the scene.

An interactive female virtual avatar was created using the ARP toolkit in order to have a highly interactive application and to create a more engaging presentation of the history of the town. Thus, the avatar was design to act as a virtual guide which responds to user's questions related to the building and events in the town. This is achieved by modeling appropiate gestures animations and adding natural language understanding to the application.

A "virtual tour" was created by generating interesting route paths in the environment. Six locations have been selected for the user to visit in the virtual reconstruction. The user navigates from one to another by clicking on labels "floating" in the sky. Once at a location, the user can look around, rotating the view or move freely with the use of keys. The user can request more information about any of the six locations in town using the following approaches: i) typing a question or ii) "pointing&clicking" on one of the predefined questions. The user also has access to a webpage when arriving at certain locations.

In this environment seen in figure 7, priority was given to building a mock-up that will allow users with different backgrounds to be part of the full interactive loop (navigation-request-processing-response-navigation) of interaction with a virtual guide. Taking into account contextual (location on the site) information about the user during the interaction provides a first impression of what natural interactive systems can achieve for navigation through Cultural Heritage sites.

Figure 7: *Avatar guide interacting with users of the virtual city of Wolfenbüttel*

The usability and acceptability of the Wolfenbuttel 3D Interactive application was investigated in order to identify the effectiveness and satisfaction of users when interacting with this type of 3D environment. In general testers found the application a good and entertaining representation of a historical place. Avatars were regarded as useful to add life to the environment, and users suggested that the levels of interactivity of the application should be higher. In addition, once testers understood the navigation and interaction mechanisms, they find the system easy to use.

Users had a more divided opinion on the best way to use these systems for presenting heritage and how they can replace current presentation systems. Testers tend to perceive them more as entertainment. This highlighted the need to make a real connection between the heritage/artefacts in display and the virtual environment. They should enhance the museum/site experience rather than try to be the experience themselves. The results of the testing are presented in [REMM*07].

5. Conclusions

The paper has presented results on the Characterize NEWTON within the EPOCH project. Both frameworks presented, vhdPLUS and UEA Scene assembly framework make use of open source generic frameworks to provide vertical middleware for constructing interactive 3D Virtual Environments. Furthermore, the paper has presented example of applications which can be built with this frameworks as a demonstrator of their potential for the Cultural Heritage sector.

The software is available from: [vhd08a] in the case of vhdPLUS and contacting the developers for the Scene Assembly Toolkit.

6. Acknowledgements

This work has been conducted as part of the EPOCH network of excellence (IST-2002-507382) within the IST (Information Society Technologies) section of the Sixth Framework Programme of the European Commission.

Our thanks go to the others involved in the realisation of this work: the research team at the University of Brighton, in particular to Michel Genereux and David Morris. Also to the resarch team at MIRALAB, EPFL and the University of East Anglia. Thanks are also due to the Graphics Research Groups at the Technical Universities of Braunschweig and Graz for their involvement in the development of the Wolfenbüttel application.

References

[CID08] CIDOC CONCEPTUAL REFERENCE MODEL:. http://cidoc.ics.forth.gr/, 2008.

[Col08] COLLADA:. http://www.collada.org/mediawiki /index.php/Main_Page, 2008.

[MHY*07] MAÏM J., HAEGLER S., YERSIN B., MUELLER P., THALMANN D., GOOL L. V.: Populating ancient pompeii with crowds of virtual romans. In *in the 8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'07)* (2007).

[MTP06] MAGNENAT-THALMANN N., PAPAGIANNAKIS G.: Virtual worlds and augmented reality in cultural heritage applications. In *In "Recording, Modeling and Visualization of Cultural Heritage"* (2006), pp. 419–430.

[NMTCY07] NADIA MAGNENAT-THALMANN ALESSANDRO FONI G. P., CADI-YAZLI N.: Real time animation and illumination in ancient roman sites. In *In "The International Journal of Virtual Reality"* (2007).

[Ope08a] OPENSCENEGRAPH:. http://www.openscenegraph.org/projects/osg, 2008.

[Ope08b] OPENSG:. http://opensg.vrsource.org, 2008.

[Pon04] PONDER M.: Component-based methodology and development framework for virtual and augmented reality systems. In *PhD Thesis* (2004).

[PPM*03] PONDER M., PAPAGIANNAKIS G., MOLET T., MAGNENAT-THALMANN N., THALMANN D.: Vhd++ development framework: Towards extendible, component based vr/ar simulation engine featuring advanced virtual character technologies. In *In "Proceedings of Computer Graphics International"* (2003), pp. 96–104.

[REMM*07] RODRIGUEZ-ECHAVARRIA K., MORRIS D., MOORE C., ARNOLD D., GLAUERT J., JENNINGS V.: Developing effective interfaces for cultural heritage 3d immersive environments. In *in the 8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'07)* (2007).

[vhd08a] VHDPLUS SOURCEFORGE REPOSITORY:. http://www.sourceforge.net/projects/vhdplus, 2008.

[vhd08b] VHDPLUS SOURCEFORGE WEBSITE:. http://vhdplus.sourceforge.net, 2008.

EPOCH Conference on Open Digital Cultural Heritage Systems (2008)
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

Interoperable multimedia mobile services for cultural heritage sites

N. Ryan[1], P. Mohr[2], D. Manzaroli[2], G. Mantovani[2], S. Bartolini[2] , A. D'Elia[2], M. Pettinari[2], L. Roffia[2], L. Sklenar[1], F. Garzotto[3], T. Salmon[2]

[1]Computing Laboratory, University of Kent, UK
[2]ARCES, Università di Bologna, Italy
[3]Department of Electronics and Information, Politecnico di Milano, Italy

Abstract

According to the ancient Romans, "Delectare, docere, movere" are the goals of eloquence. To be accepted by museums, landscapes and archaeological sites, technology has to win the same challenge. Is technology unobtrusive enough to avoid compromising the emotional involvement that makes a visit to a cultural site unforgettable? Can it achieve a dissemination of the information in such a way that it is understood better? And how can technology be used to increase visibility and understanding of the numerous sites that are not yet able to attract the amount of people they deserve?
This paper presents the authors' vision on these questions, reporting on the activities carried out by the "mobile and ambient systems" work group of EPOCH as part of the CIMAD project.
A central part of CIMAD is the creation of services for visitors and archaeological sites as well as making parts of the overall vision a reality. The CIMAD services are based around the MobiComp context infrastructure, enabling the services to exchange context information and information to be displayed to the user. As the EPOCH network is beginning to dissolve we will discuss possible next steps, associated risks and opportunities of continuing this project.

Categories and Subject Descriptors (according to ACM CCS): H.3.5 [Online Information Services]: Mobile Guides, Museum and CIMAD

1. Introduction

Cultural Heritage (CH) is becoming a greater attraction factor for tourism worldwide and many countries are aiming to offer lower-cost, but higher-quality content and services that can provide better visibility and improved understanding to their museums, sites and landscapes.

In order to improve CH perception in a cost-effective way, many actors should be involved, i.e. tour operators, service providers, CH sites owners, telecom operators and technology providers, and many integrated services should be provided. The role of technology is to provide tools to enable service implementation and deployment. Services should address Site management, Data Collection and Content Delivery.

A structured approach to the use of technology in CH is required and EPOCH has taken on this challange of "overcoming fragmentation in open cultural heritage" and has devised a reference architecture for all CH related taskes from data collection to data presentation. The architecture is discussed in another section of this issue and its abstract view is shown figure 1.

This paper concentrates on the mobile user's view of the architecture, where mobile users can be researchers, visitors and museum staff. Their context is characterized by their position, activity, profile, preferences and by the devices they use to interact with both physical and digital CH.

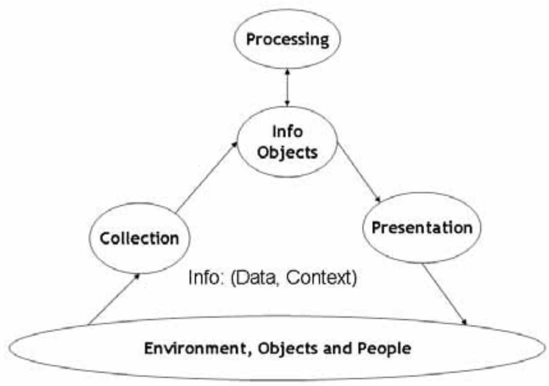


Figure 1: *EPOCH architecture.*

Mobile users need services integrated within the environment and widely accessible, requiring them to be available on a wide range of user equipment. Ideally, they should be automatically customized to the user's profile and preferences, and they should be smoothly and selectively accessible wherever they are needed.

The key to this accessibility and interoperability are seamless connectivity and seamless services; requiring platform interoperability (same services run on different platforms), data interoperability (services work on a common representation of data, and are independent from their data sources) and we need also network interoperability, to be always connected to the best available network, as well as session interoperability, to support seamless working session migration.

Figure 2 shows how this vision is approached by EPOCH: in order to separate services from their heterogeneous data sources EPOCH has adopted an ontology to describe cultural heritage objects and a common representation of context. CH data objects, coming from a

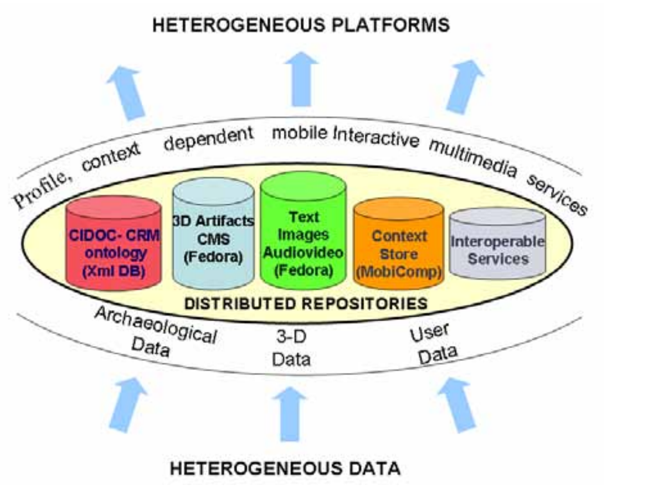


Figure 2: *Interoperability for service integration: the mobile user perspective.*

wide range of sources, e.g. from archaeological sites or from intermediate interpretation and processing steps such as 3-D modeling, are given semantic meaning by a CH ontology and are stored in a *content* management system. Context data, no matter where it originated from, e.g. coming from sensors in the environment or from direct user input, are wrapped into context elements to achieve a common format — they are managed by a *context* management system. The most suitable services should be offered to the users based on their current context.

The proposed approach is not a "Near Market" view, since the picture outlined is expected to become a reality in the second half of the next decade, as it builds on technologies, systems and tools that are envisaged by the European 7th Framework Program whose time-span extends to 2013.

The goal of this paper is to demonstrate the potential of the proposed approach with a prototype framework based on current technologies — which are mainly in a development state. A prototype framework, called CIMAD (standing for "Common Infrastructure/Context Influenced Mobile Acquisition and Delivery of cultural heritage data"), supports demonstration level services which can simplify CH management and exposure. Services range from site management to visitor management and include, e.g. registration services, visitor flow monitoring, visitor guiding with support for user orientation and content delivery. The range of supported services could be extended for example with tools for site survey and data acquisition.

CIMAD is meant to seamlessly support heterogeneous device types with different form factors and usage models, i.e. custom made devices hired on site, privately owned PDAs and smart phones, and different location technologies, e.g. WiFi based and GPS based ones. Even if necessarily incomplete and preliminary in many respects, CIMAD demonstrates a future application development scenario, supporting the integration of services applicable to many CH environments. In order to provide such services, CIMAD builds on top of two infrastructure components supported by EPOCH: a context management system called MobiComp [Rya05] and a content management system built on top of the Fedora digital repository [UoV]. A work in progress demonstration of CIMAD services, named "Smart Museums, Sites and Landscapes — From Visitor Guides to Collection Monitoring" was setup and tested during a number of EPOCH sponsored events across Europe. This research is being carried out by an international research team which emerged from within the EPOCH framework.

In the section 2 the CIMAD framework is described, followed by the description of MobiComp in section 3 and Fedora in section 4. The implemented services are described in section 5 and a conclusion is presented in section 6.

2. CIMAD Framework

A wide range of CH applications have been developed in the past, mostly isolated and developed from scratch, not very flexible and hardly reusable. However, in the field of context-awareness an increasing number of applications are being developed out of modular re-usable building blocks which are combined through a context infrastructure [DSA01, Win01, CR02]. Based on such a context infrastructure, CIMAD aims to introduce the same modularity to the area of CH applications, by providing a common framework that allows overlapping functionalities and context elements to be re-used. The goal of CIMAD is to speed up and simplify the development process of CH applications — aimed both at developers and field experts with limited IT knowledge. The most widely used functionalities in CH applications are:

- Dynamic adaptation of content, for example to device characteristics, user preferences and profile.
- Seamless data acquisition in fieldworks, for example with contextualization of notes and pictures.
- User context detection from sensors, with a focus on position detection followed by guidance.
- Context abstraction for detecting meaningful user states, for example walking or looking at a particular exhibit.

Through an extendable basis of modules for the most widely used functions, the overall goal of the common framework is enabling a wide number of developers, with different levels of experience, to efficiently develop modular context-aware multi-channel CH applications which are interoperable. The modules developed by different developers form the asset of CIMAD, speeding up the development process of new applications through re-usability of existing modules. Aiming to cater for the different levels of users ranging from archaeologists, museum curators to experienced developers is one of the biggest challenges of CIMAD. This challenge is confronted through a flexible structure, providing support and guidance at different levels — for further details please refer to [RMR*07].

One of the main applications that can be set up within CIMAD is a visitor guide. As an example, the implementation process of a CIMAD interactive multimedia guide could look like the following:

- Cultural heritage specialists, i.e. museum curators or site experts, prepare the multimedia content and select the appropriate user interface.
- Curators prepare a "map component" associating each exhibit to its "context", e.g. physical location.
- Curators identify the criteria for organizing multimedia content into "thematic" or "geographic" tours.
- The site management team together with the developers select the desired devices and technologies for delivering the guided visits, i.e. PDAs and location technology used. Based on the selected devices and technologies the developers construct the visitor guide.

3. Context Management: MobiComp

MobiComp [Rya05] is a context management infrastructure tailored to the needs of CH. Its core element is the ContextService, acting as a store for context information and enabling coordination between the components of context-aware applications. The storage components behind the ContextService interface can be configured to support different scales of context-aware applications: simple standalone applications, multiple applications on a single device and applications spanning multiple devices. Three components exist for interacting with MobiComp: trackers, listeners and aggregators. A tracker is a MobiComp component that acts as a context producer. Trackers register their availability and capabilities by sending appropriate information to the ContextService. Their purpose is to collect raw context data from sensors, such as GPS receivers, and other dynamic or static sources, including configuration files for device capabilities and user-preferences. Trackers transform their input into context elements which are then put into in a database. Therefore, applications that need context information retrieve it directly from the MobiComp server and do not have to care about the actual context sources.

A listener is a MobiComp component that receives notification of ContextEvents from the ContextService and performs some action based on the context element carried by the event object. They receive event notifications whenever a context element is put into or removed from the store. On receiving a notification, the listener may get the element from the store and use it as required.

An aggregator is a MobiComp component that combines the behaviour of both a tracker and a listener. Aggregators monitor events from the ContextService, rather than a sensor device, and apply a transformation before returning a new element to the database. Aggregators can combine several low-level sensor elements to produce an element at a higher level of abstraction. For example, temperature, door, window and light sensor information might be used to determine room occupancy. Other aggregators may perform simple transformation services, i.e. converting latitude and longitude coordinates from a GPS sensor to coordinates on an appropriate local or national grid. Many non-trivial context-aware applications utilise a number of complex context aggregators, e.g. the FieldMap application described in [vLR01]. To ease communication between infrastructure components, context elements are represented in the form of a XML document based on ConteXtML [Rya05]. The elements carry a production timestamp, a default validity period, and a privacy level indicating how they may be disseminated through the ContextService.

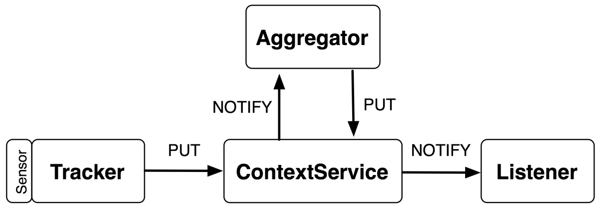


Figure 3: *MobiComp infrastructure.*

Through the above mentioned components the desired interoperability described in the introduction is achieved for context information — how it can be achieved for descriptions and multimedia related to artefacts and exhibits is described in the next section.

4. Fedora Content Store And The Content Adaptation Layer

Fedora [UoV] is the content repository system adopted by EPOCH. In addition to the content repository, Fedora provides a collection of tools and interfaces for creating, managing, and disseminating "Fedora digital objects" (FDO) stored within the repository. A FDO allows the original format of an object to be stored, along with metadata, i.e. in the format of the Dublin Core [DCM]. Through format adaptation components it is possible to perform a format conversion of an FDO in real-time, allowing requests from external applications for a specific format to be satisfied, e.g. HTML, PDF and JPEG. For example, a scaled-down or a greyscale version of an image can be retrieved according to the device characteristics. The requests and retrievals are performed through standardised interfaces based on REST and SOAP — within the CIMAD architecture the "Content Adaptation Layer" (CAL) is used.

Fedora enables the multi-channel paradigm, by allowing FDOs which were produced once to be adapted at run time according to the user and device context. In figure 4, the interaction of an application with Fedora is shown. The application publishes context on MobiComp through trackers and the CAL retrieves it through listeners. Once Fedora is aware of the context, it can provide the application with content which has been adapted to best fit that context.

5. CIMAD Services

As part of the Interactive Salon, an EPOCH event, a number of CIMAD prototype services supporting visitors and staff have been demonstrated. Some of the services run stand alone on "any" PDA or Smartphone, e.g. Symbian and Windows mobile ones, equipped with a web browser. These services demonstrate the extent of interoperability which can be achieved across platforms. Other services run on a server and can be invoked by any client application, e.g. through web pages — an example is the Path Finder, a service to provide the best path between two points on a site. Other services are based around wearable sensor kits, e.g. inertial sensors and RFID readers — not yet available on most PDAs and Smartphones. They extend the functionality of available PDAs and smart phones, and demonstrate the capabilities of future mobile devices. The additional functionality provided by the sensor kits currently consist of indoor positioning and tracking. These services are interesting because they

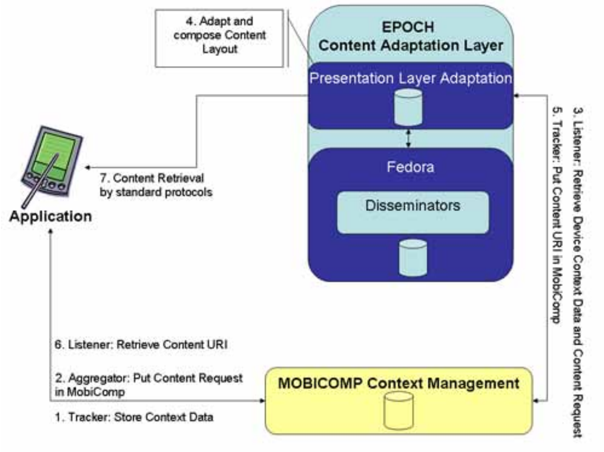


Figure 4: *Fedora content store and CIMAD integration.*

demonstrate that with insignificant impact on the hosting environment, i.e. a small number of RFID tags spread in the environment, they increase the effectiveness of an application.

Several different implementations of visitor guides based on various CIMAD components were proposed within EPOCH — the collection of guides can be seen in figure 5. Next the main services are described individually; for further information and a description of the individual components used please refer to [RMR*07].

Figure 5: *All guides Guide.*

- One of the guides, realised on a PDA and based on IR beacon location, displays information about the exhibit the visitor is standing next to. The MobiComp components used for this guide are for example an IR beacon tracker and an URL display listener, in addition to an aggregator able to convert the IR beacon stream to a sequence of URLs, forming the "virtual path" of the visitor through the museum. The IR technology of this guide can be exchanged with RFID to create a guide which does not require line of sight.
- Another guide is based on computer vision components developed by colleagues at ETH, Zurich. This is a Museum Guide [BFvG05, RMR*07] implemented on a Tablet PC using a conventional USB webcam to acquire images and to present users with a viewfinder window. When a user wants to get information about an exhibit, the information button on the tablet needs to be pressed when the exhibit is within the viewfinder window. If the exhibit is recognized using an image recognition algorithm as depicted in figure 6, corresponding exhibit information is displayed.

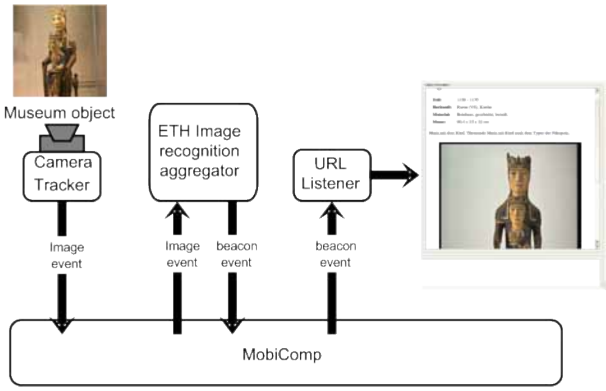


Figure 6: *Camera tracker Guide.*

- The semacode based guide can be run on standard mobile phones with a built in camera and a GPRS connection — the guide is installed on the phone via a "Bluetooth Kiosk". The visitors can take pictures of semacodes [Sem] situated next to exhibits they are interested in. Instead of an image recognizer like in the above guide, a semacode recognition aggregator finds and decodes the semacode tag in the captured image. This external module uses the Semacode.org software development kit [Sem]. Based upon the url decoded, the corresponding exhibit information is displayed on the screen of the mobile phone. The similarities between the guides highlight that the modularity of the MobiComp approach is powerful and allows for components to be re-used and bundled together quickly and easily under the CIMAD architecture.

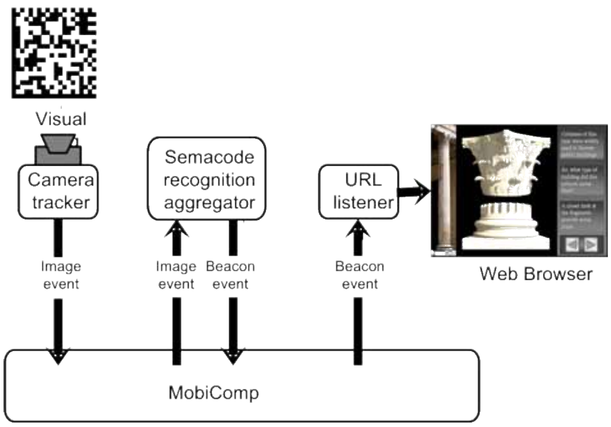


Figure 7: *Semacode Visitor Guide.*

- A visitor guide for the archaeological site of Palmyra in Syria is implemented on a standard PDA. The application is based on a CIMAD map component that listens to "location change" events and reacts updating the visual map —see figure 8— of the site and

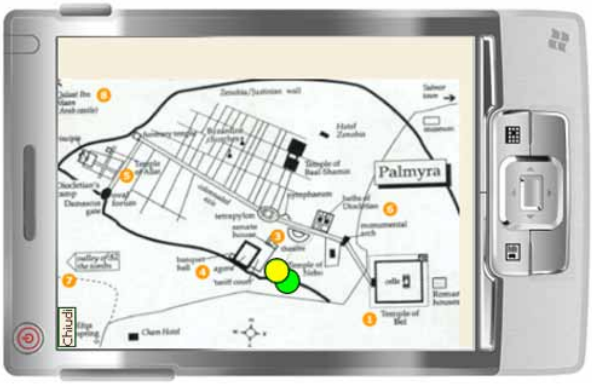


Figure 8: *Palmyra Visitor Guide: Map.*

presenting a location dependant interface to the user once a Point-Of-Interest is reached — see figure 9. The MobiComp components are a GPS tracker and a location listener. Content is adapted to the requirements of the PDAs through Fedora and the CAL.

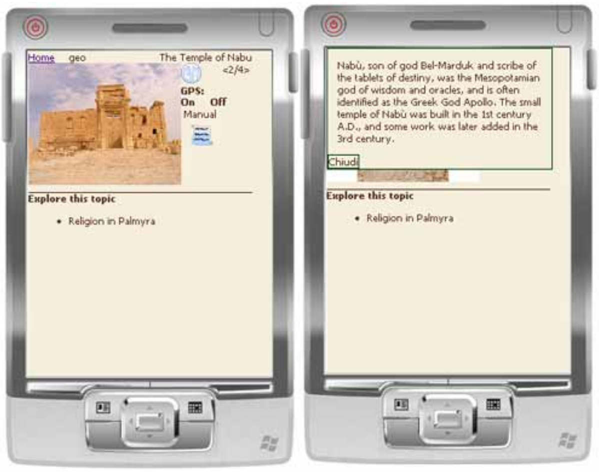


Figure 9: *Palmyra Visitor Guide: Content Presentation.*

- Yet another guide is the "Content navigator", which a user can run solely on a device with a web browser and connectivity — not requiring any other user hardware or software. When the Content navigator is loaded on the user's web browser, the browser parameters, i.e. the screen resolution, are determined automatically and are send to the server based content repository together with user preferences entered on the webpage. Once this data is received by the server, a disseminator is activated and the right content is provided to the user in the right format for the user's device and displayed through the web browser. The content can be adapted by Fedora, where it is stored as FDOs.
- The inertial tracking guide does not require any user interaction, it automatically displays information about the exhibit the user is closest and facing to. User position and orientation are provided by an inertial location tracker strapped around the users waste — see figure 10. The tracker provides an estimate of the current user position, with respect to a pre defined starting point. Any type of beacon, ranging from an object recognizable by a camera to an RFID tag can be used to set the tracker's initial position. If RFID tags are used, the tags are distributed around the site, providing the possibility to reposition the tracker when a significant position error has been accumulated — inherent in the inertial tracking system.

Figure 10: *Inertial tracking guide.*

- All of the above mentioned guides may benefit from the Path-Finder, a service providing the shortest path between a source and a destination point. The destination point may be chosen by the user, or it could be automatically recommended by a thematic tour service. The starting point can also be set by the user, or it could for example be the current location of a nearby exhibit, recognized through an RFID tag. For an example screen shot please see figure 11. The guide can be installed on a Web server or directly on the user's device. In order to function in a given museum or site, a map is required. This map can, for example be scanned by the curator and easily converted into the required graph format by the "Path Finder Remapping Utility", which runs on any desktop computer and comes with a demo application — enabling the graph creation by non experts. The "shortest" path is searched on the graph with an A* algorithm [Pea84]. Beside providing the shortest path between two conspicuous points, the "Path Finder" may also help the visitors of cultural heritage sites to find their way when looking for a specific exhibit, an emergency exit, the book shop, or any other location within a museum or site, even if they do not explicitely know their position. To this end a tool to estimate the user current position is needed. An ideal combination is the inertial tracking guide and the path finder; by dynamically evaluating the shortest path between the current position and the destination, the user can continuously be guided to a predefined target, very much like GPS navigation systems. This novel type of guide, that we could call ENDD, Epoch Navigation and Discovery Device is currently being tested at the Museo Civico Archeologico in Bologna, that kindly agreed to host the test within an informal co-operation with EPOCH's mobile and ambient systems group. For performance and scalability reasons ENDD is currently using a version of the Path Finder installed on the user's device.

In order to demonstrate that the CIMAD architecture can be used horizontally, supporting services for all mobile museum actors, i.e. not only visitors but also staff and management, two prototype man-

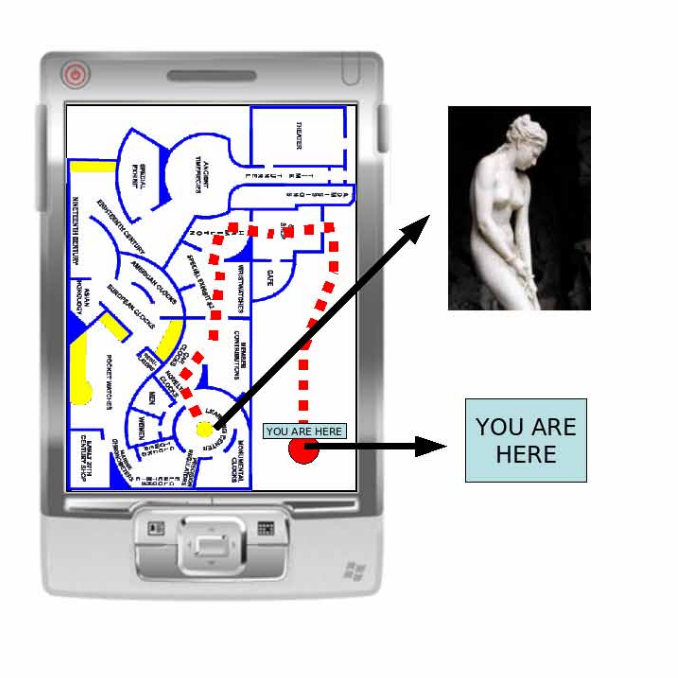


Figure 11: *Path Finder: after clicking on an exhibit icon, the system calculates the shortest path from the user's location to the selected exhibit.*

agement services were demonstrated and deployed during the Interactive Salon. One is a museum presence monitor which keeps track of the number of visitors currently in the exhibition and of the overall total number of visitors. The monitor application is notified by the visitor location tracker when a visitor enters or leaves the exhibition.

The other one is a meuseum registration desk service, which registers visitors who wish to use a context aware guide with the CIMAD services. The visitor's details (profile and preferences) are entered and the chosen guide is selected. This information is made available to all applications through the ContextStore and also triggers the configuration of the chosen guide for this particular visitor. Visitors can remain anonymous or can sign up for post-visit online services. For further information regarding implementation and details about the individual CIMAD building blocks used in the services please refere to [RMR*07] and [RRM*06].

6. Conclusions

This paper describes the mobile and ambient systems group's vision of the EPOCH mission "Overcoming Fragmentation in open cultural heritage". Starting from the consideration that increasing visibility and understanding of Cultural Heritage sites is an usatisfied social need [Ant07], this paper focused on how mobile technology could contribute. No isolated, individual solution with a large number of functions is desirable, while an infrastructure supporting seamless connectivity between interoperable contents is needed. Also a new interaction model is envisaged, where the users define their goals, e.g. "guide me to that particular exhibit and tell me if there is something interesting for me along the way", rather than select a function, i.e. "show me the map" or "show me the list of exhibits". Eventually this approach requires that the objects in the environment interact with each other, providing services and multimedia content to anybody. Objects equipped with identification devices and filled with multimedia content will make the environment user friendly and interactive. Research needs to solve performance and system issues in order to meet interoperability, scalability and privacy requirements. These are considered primary issues both in the long term and mid-term european reserch agenda, respectively within 7FP and related initiatives. Cultural heritage will keep its primary role in offering test beds to new technologies, new systems and new services, and EPOCH, as a 6FP NoE had the opportunity to anticipate similar research currently in the agenda of other priority domains. CIMAD —a prototype framework for interoperable context-aware applications— was described in this paper, together with its components, and with several examples of services for museums and archaeological sites.

Lessons were learned and issues requiring further research were identified. For example, it was found that occasionally, total separation between applications and data sources, as depicted in 2, cannot be achieved with current architectures and performance levels. There is also the need to further evaluate the proposed solutions, as well as to define a strategy to consolidate the EPOCH vision. EPOCH is coming to an end with its final event in Roma, which is a unique occasion to demonstrate the potential impact of the EPOCH architecture together with the associated models of interaction with cultural heritage. However, the event should celebrate the beginning of a new research cycle, not its end.

7. Acknowledgements

We would like to thank the EU and EPOCH for supporting this work, and the various museums and conferences for making concept verification in a real world environment possible. EPOCH, the European Research Network of Excellence in Processing Open Cultural Heritage, is funded by the European Commission under the 6th Framework Programme, Contract no. IST-2002-507382. However, the content here reflects only the authors' views and the Commission is not liable for any use that may be made of the information contained herein.

References

[Ant07] ANTINUCCI F.: *MUSEI VIRTUALI*. LATERZA, 2007.

[BFvG05] BAY H., FASEL B., VAN GOOL L.: Interactive museum guide. In *Smart Environments and their Applications to Cultural Heritage* (September 2005), Ryan N., Cinotti T. S., Raffa G., (Eds.), EPOCH Publication, Archaeolingua, Budapest, Hungary.

[CR02] COUTAZ J., REY G.: Foundations for a theory of contextors. In *Proc. CADUI02* (2002).

[DCM] DCMI: Dublin core metadata initiative. `http://dublincore.org` (accessed Feb.2008).

[DSA01] DEY A., SALBER D., ABOWD G.: A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-Computer Interaction 16* (2001).

[Pea84] PEARL J.: *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. Addison-Wesley, 1984.

[RMR*07] RAFFA G., MOHR P. H., RYAN N., MANZAROLI D., PETTINARI M., ROFFIA L., BARTOLINI S., SKLENAR L., GARZOTTO F., PAOLINI P., CINOTTI T. S.: Cimad - a framework for the development of context-aware and multi-channel cultural heritage services. In *Proc. International Cultural Heritage Informatics Meeting (ICHIM07)* (Toronto, Canada, October 2007).

[RRM*06] RYAN N. S., RAFFA G., MOHR P. H., MANZAROLI D., ROFFIA L., PETTINARI M., SKLENAR L., STEFANO L. D., CINOTTI T. S.: A smart museum installation in the stadsmuseum in stockholm - from visitor guides to museum management. In *EPOCH Workshop on the Integration of Location Based Systems in Tourism and Cultural Heritage* (Brussels, Belgium, November 2006).

[Rya05] RYAN N.: Smart environments for cultural heritage. In *Reading the Historical Spatial Information in the World, Proc. 24th International Symposium* (International Center for Japanese Studies, Kyoto, Japan, February 2005).

[Sem] SEMACODE: Semacode software development kit. http://www.semacode.org (accessed Feb.2008).

[UoV] UNIVERSITY C., OF VIRGINIA U.: Fedora digital repository system. http://www.fedora.info (accessed Feb.2008).

[vLR01] VAN LEUSEN M., RYAN N.: Educating the digital fieldwork assistant. In *Proc. of Computer Applications and Quantitive Methods in Archeology Conference* (Gotland, 2001).

[Win01] WINOGRAD T.: Architectures for context. *Human-Computer Interaction 16* (2001), 401–419.

EPOCH Conference on Open Digital Cultural Heritage Systems (2008), pp. 1–7
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

Encoding Cultural Heritage Information for the Semantic Web. Procedures for Data Integration through CIDOC-CRM Mapping

Ø. Eide[1], A. Felicetti[2], C.E. Ore[1], A. D'Andrea[3] and J. Holmen[1]

[1]Unit for Digital Documentation, University of Oslo, Norway [2]PIN, University of Florence, Italy [3]CISA, University of Naples "L'Orientale", Italy

Abstract

This paper describes the background and methods for the production of CIDOC-CRM compliant data sets from diverse collections of source data. The construction of such data sets is based on data in column format, typically exported for databases, as well as free text, typically created through scanning and OCR processing or transcription.

Categories and Subject Descriptors (according to ACM CCS): H.3.1 [Content Analysis and Indexing]:

1. Introduction

As part of the EPOCH Network of Excellence, several tool chains for cultural heritage have been investigated, and a complete framework for the mapping and management of cultural heritage information in a semantic web context has been developed. PIN (University of Florence, Italy), CISA (University of Naples "L'Orientale", Italy) and EDD (University of Oslo, Norway) have been central within the EPOCH context in these activities through the AMA project.

AMA (Archive Mapper for Archaeology) is one of the NEWTON projects (NEW TOols Needed). EPOCH's NEWTONs aim at plugging gaps in the digital processing of Cultural Heritage on the basis of existing tools or tools under development for Epoch's Common Infrastructure (WP 3.3). One of the main goals is to develop entire scenarios that can handle multiple steps in such processing chains.

2. General Mapping problems in cultural heritage

Access to data, interoperability and standard structure are vital to guarantee the usability and usefulness of Cultural Heritage records, which are digitally archived in a myriad of different ways and are stored separately from each other. In many ways, they are as dispersed as the material culture they refer to.

Proposed standards have not yet overcome the diffidence of culture professionals and heritage policy makers because of the difficulty of mapping existing data structures to them. Mapping requires skills and knowledge which are uncommon among the cultural heritage professionals with the most thorough knowledge of the material to be mapped. It is not only the absence of facilitating tools, but also the existence of practices, legal obligations and the lack of a clear motivation that has as yet delayed or reduced the creation of such mappings to a handful cases.

There have also existed a hope that automatic tools, such as the ones developed in natural language processing research, will solve the integration problems. Although such tools have been and will continue to be helpful in many areas, solutions to the data integration problems have to be solved at a different level, and human interaction will continue to be necessary.

3. The AMA project

The aim of the AMA project was to develop tools for semi-automated mapping of cultural heritage data to CIDOC-CRM, ISO 21127 [CDG*05]. The reason for this investment was that such tools would enhance interoperability among the different archives and datasets produced in the field of Cultural Heritage [D'A06c]. For this purpose we created a set of tools able to extract and encode legacy information coming from diverse sources, to store and manage this information using a semantic enabled container and to make it available for query and reuse.

The tool set developed in the AMA project includes:

- A powerful mapping application for the creation of mappings from existing datasets
- A tool for mapping cultural heritage information contained in free text into a CIDOC-CRM compliant data model
- Templates describing relations between the structure of existing archives and CIDOC-CRM
- A semantic framework to store, manage and browse the encoded information providing user-friendly interfaces

All tools have been developed as part of the EPOCH and AMA projects and can be freely downloaded from the EPOCH web site with no usage restrictions. The rest of this article describes the background for the tools as well as their implementation and use.

4. Data standards

4.1. CIDOC-CRM

International standards and ontologies for data encoding are crucial to speed up interoperability and the process of integration. CIDOC-CRM, which is created to capture the richness typical of Cultural Heritage information, fully fits our needs: its classes and properties work perfectly to capture the concepts underlying database structures, providing a high level of data integration.

CIDOC-CRM is a formal ontology intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information. It was developed by interdisciplinary teams of experts, coming from fields such as computer science, archaeology, museum documentation, history of arts, natural history, library science, physics and philosophy, under the aegis of the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM). The harmonisation of CIDOC-CRM and IFLA's FRBR [FRB98] is completed and have been submitted to IFLA for comments. The EAD has already been mapped to CIDOC-CRM [TD01]. This shows that CIDOC-CRM falls well in with a larger memory institutional framework.

CIDOC-CRM is defined in an object oriented formalism which allow for a compact definition with abstraction and generalisation. The model is event centric, that is, actors, places and objects are connected via events. CIDOC-CRM is a core ontology in the sense that the model does not have classes for all particulars like for example the Art and Architecture Thesaurus with thousands of concepts. CIDOC-CRM has little more than 80 classes and 130 properties.

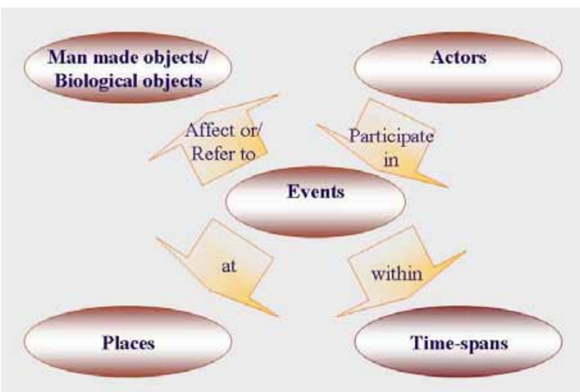


Figure 1: *The most central classes and properties for data interchange in CIDOC-CRM.*

4.2. TEI

TEI is a consortium of institutions and individuals from all over the world. The TEI is also a set of guidelines for the encoding of textual material, and it is a set of computer readable files. The guidelines and the computer readable files specify a set of rules documents have to adhere to in order to be accepted as TEI documents.

One of the main goals of TEI is to capture a wide range of intellectual work. It is used extensively in areas such as edition philology, but it is quite possible to create a detailed encoding scheme for e.g. archaeological grey documents and include it as a local extension of TEI, creating documents very different from most other TEI documents [EH06].

According to the TEI guidelines [TEI08, section iv.i], there are three primary functions of the guidelines:

- guidance for individual or local practice in text creation and data capture;
- support of data interchange;
- support of application-independent local processing.

TEI is also an important discussion forum, or a set of fora, both through meetings and on-line media, as well as through articles and books. Few encoding problems will be faced that has not been discussed somewhere in the guidelines or in the literature surrounding TEI.

5. From database to CIDOC-CRM

The AMA project was aimed at overcoming the mapping problems through the implementation of a flexible mapping tool in order to facilitate the mapping of different archaeological and museum collection data models (with various structured, as well as non-structured data, i.e. text description) to a common standard based on CIDOC-CRM [D'A06b]. The necessary information was extracted from such mappings to convert individual datasets to a common, CIDOC-CRM compliant structure [DMZ06].

5.1. AMA Mapping Tool

The design of the tool design had to start from an accurate knowledge of the material it handles, that is, the source data structure or, at least, their common matrix as generated by national regulations and most common practices. It also had to be tested on a wide sample of such material.

In principle, the tool had to operate on any preliminary, fragmentary or old archaeological format or museum data (both structured and non-structured), and modern datasets, mapping their structure into a single, standardised CRM-compliant system, easily accessible for investigation by means of a web-based interface. Since CIDOC-CRM deals well with museum collections, while work still had to be undertaken for archaeological excavation data, monuments investigation and landscape analysis, the project also considered these aspects.

The tool is based on the concept of "template", this being the instantiation of the abstract mapping between the source data structure and the mapping target. A template documents the source data structure and the mapping, using formal XML documents, and automatically supports the conversion of actual datasets when desired. Templates capture the semantic structure, as well as the intellectual content, of the sources and their transformations. They ensure the modularity and future maintenance of the system, in case new or extended versions of the standard are available [D'A06a].

The AMA Mapping tool developed at PIN University of Florence and CISA University of Naples comes with an on-line interface written in PHP through which the mapping template can be defined and exported for further use. Features include the possibility to upload the XML starting schema (i.e. the XML description of the user's databases or an existing ontology), to perform the mapping operation using the simple and intuitive on-line interface and to export the high level mapping files to be used in further data conversion and extraction. The tool also provides the possibility to upload a different target ontology (CIDOC-CRM is the predefined one) and an advanced set of mapping features for complex mapping definitions, such as for the creation of new entities and properties to represent implicit elements and relations, the creation of shortcuts to simplify the mapping process and a graphic visualisation of simple and complex relations obtained during the mapping process.

High level templates coming out of AMA can be directly imported and used by the MAD framework on information stored in legacy databases in order to get perfectly mapped semantic archives, as described below [AMA].

6. From text to CIDOC-CRM

The objectives of the AMA NEWTON is to create tools to assist the construction of mappings from archaeological archive material, reports, catalogues and databases to CIDOC-CRM. An important step in this process is the extraction of information from free text documents into a well defined format with a predefined mapping to the CIDOC-CRM or to a subset thereof.

There are several strategies to perform such extraction of information from running texts:

1. manual extraction by reading and keying the information of interest into a predefined form
2. encoding the information in the actual text by the use of XML
3. using statistical based topic and summary extracting software

The first strategy has been and still is the most widespread. To have a specialist reading and keying the core data into a registration form seems faster measured by the number of pages processed per day than encoding the content of the text by the use of XML. However, the job requires a scholar/specialist and it is very hard to do the proof reading and virtually impossible to check the correctness of the data at a later stage. It will also be too costly to repeat this process with another focus [JHOOng].

The XML-encoding process is time-consuming but can be done by non-specialists working under the supervision of domain experts. It is also easy to track from which part of the original textual documents the information has been taken. Based on the XML-markup it is also possible to create mappings to different database formats, store the text in a free text retrieval system, or convert it to HTML or PDF for publication.

The third method is best suited for finding documents in large document collections and not for identifying highly granulated facts. It is however an interesting field of research.

The Unit for Digital Documentation (EDD) has 14 years of experience with large scale XML (previously SGML) content markup of archaeological texts as well as other categories of text [HOE04]. This is very time consuming and costly to do manually. To speed up the process many simple markup programs have been written for ad hoc use. However, it is hard to write a program that is capable to do non trivial tagging correctly. This is even more difficult than writing program for automatic topic indexing of large document collections, as the latter programs only gives a statistically based indication of the content of the documents. It may be impossible to create software to fully automatise the process; it is certainly out of scope for the AMA project to try to write such software. In the AMA project we are developing a semiautomatic tool based on well known language technological methods for speeding up and improving the encoding process.

6.1. The work process

An XML document conforming with the recommendation of TEI, extended with the necessary elements for marking up archaeologically interesting content, would be the ultimate goal for the mark up of archaeological documents and reports of which large collections exist in many countries. The encoding process can be divided in the following steps:

1. Create an electronic text from the original documents, if necessary — in some cases, the text already exist in digital form
2. Supply the electronic text with a TEI-header with bibliographical information about the original text and the current electronic text (obligatory)
3. Give the text a structural mark up i.e. identify and encode chapters, paragraphs, pagination etc. (optional and ad libitum)
4. Identify and mark up the archaeological information, that is, persons, places, excavations and other events, artifacts etc. The mark up should be done according to a predefined XML grammar designed for type of text in question. Such a grammar, also called a tag set, can be expressed as e.g. a DTD or an XML schema. This could be the extended TEI described above. In any case, the XML grammar must be based on a data model with a mapping to CIDOC-CRM or a subset thereof. "Based on" means that the tags in the text can be used as information anchors from which one can define an extraction from the text into e.g. a database defined according to the data model.

The data model described in step 4 is similar to the template described in section 5.1. The data model is mapped to CIDOC CRM and this serves as the abstract mapping between the source data structure, i.e. the predefined XML grammar, and the mapping target.

The steps 1 and 2 is not in the scope of the AMA text encoding tool. Step 3 may be in the scope, but is not central. Step 4 is the most intellectually demanding process and definitely the most time consuming part. The AMA text tool assist the encoder to speed up this process. In our large scale digitisation projects in Norway the actual mark up work has been done by unskilled low cost persons on work retraining schemas. The mark up has later been refined by using an ad hoc KWIC concordance program and other ad hoc text manipulation tools written in Perl 5. The functionality described below is based on many years experience in using such ad hoc tools.

6.2. The functionality of the AMA text tool

The tool is based on techniques from computational linguistics. A Key Word In Context (KWIC) concordance is a concordance where there is one line per hit and the hits are listed with a piece of the context at the actual hit point. Such a list may look like this (more context is usually included, but the lines are shortened to fit the column width of this paper):

of the second	angel	, these verses composed in
scroll of the	angel	on the left side of the
relating the	angel	statue to the tomb of Henry
backside of the	angels	confirm that they originally

The main feature of the AMA text tool is a KWIC concordance tool directly connected to the electronic text(s). The KWIC concordance tool is used to find a word, a phrase or a pattern of words and possibly XML-markup already in the text. The user can then study

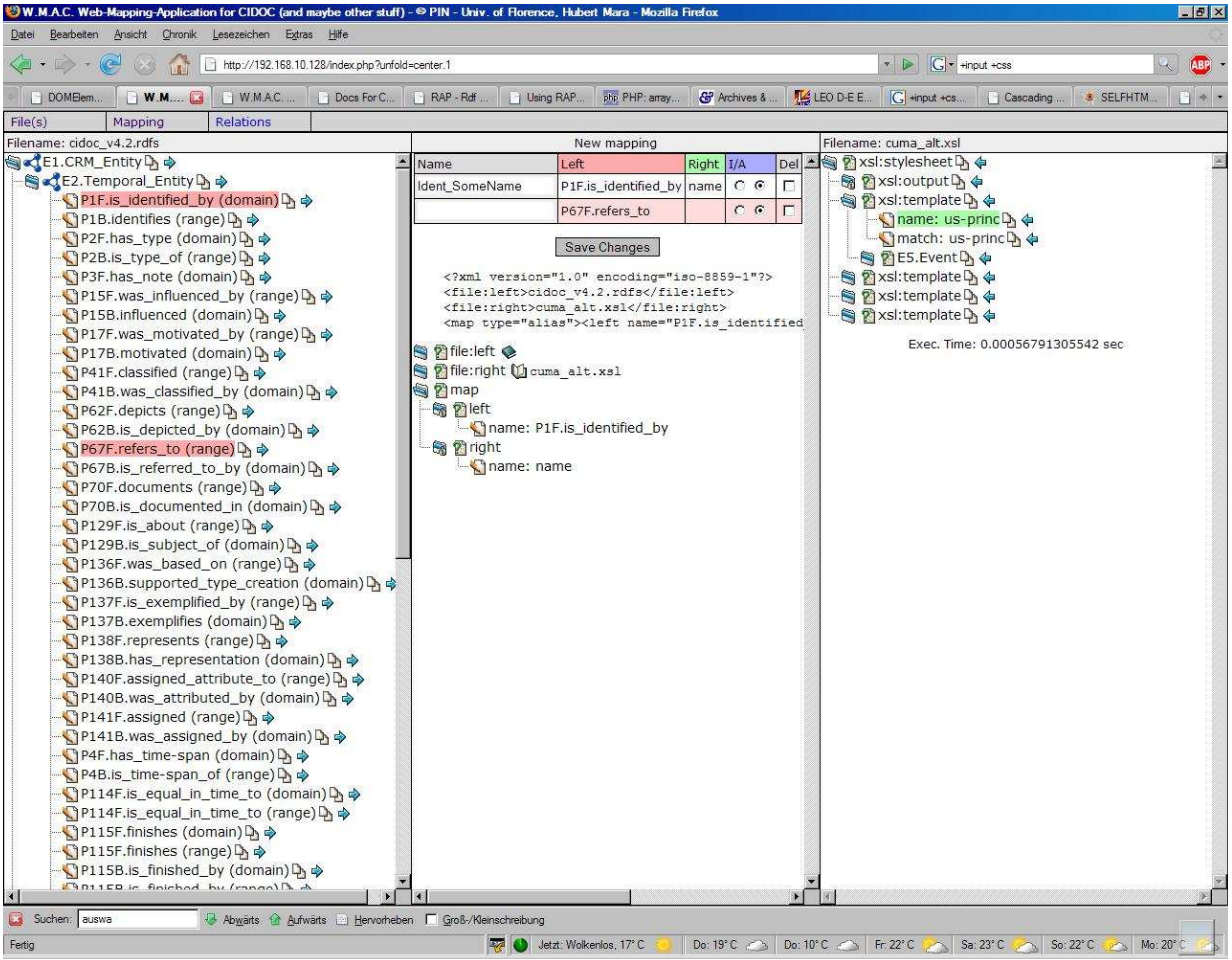


Figure 2: *A snapshot of the new AMA online mapping tool.*

the text and mark or tick which occurrences in the KWIC concordance he/she want to give a certain mark up (tag). The system then inserts the mark up in the text file(s).

This semi-automatic “search and replace” feature is implemented in a way that makes it possible for the user to create and include new algorithms both for text retrieval and for text mark up. In this way the tool features can be widened by the users when needed, more complex language specific algorithms can be added and new features of electronic text processing can be integrated.

To make the system more user friendly, the ability to parse and use an XML grammar will be included in a later version. Then it can suggest for the user which elements are allowed to be inserted on a certain point in the text. An XML grammar checker (a validator) will also be connected for use from within the program in a future version.

Another development will be to include the possibility of displaying two or several KWIC concordances. This is very useful when one need to interconnect elements by the use of ’id’ and ’idref’ attributes.

The program is written in Java and is free software, available for anyone from the EPOCH web site.

7. Storage and retrieval of mapped data

7.1. MAD

MAD (Managing Archaeological Data) is a framework originally designed to manage structured and unstructured archaeological excavation datasets encoded using XML syntax, including free text documents marked up in XML, as described above. During the development process we added more features transforming the first release into a multipurpose engine able to store and manage ontology encoded information, i.e. data structured in CIDOC-CRM compliant form. The framework can be used to browse and query such data in many powerful ways and to transform and supply semantic data on demand. The whole framework is developed using Open Source software and standard XML and W3C technology.

The first release of the tool was able to manage XML documents in a native way. It was built around eXist XML database [eXi], a Java tool easy to integrate into XML processing applications. It is able to store and index XML documents in a file-system-like structure of folders and sub-folders called collections. This allows a very simple mechanism to organise and browse XML content. The system is able to index structured documents, i.e. XML documents with a database-like structure, and also unstructured ones in the form of tagged free texts such as excavation diaries. This includes XML documents created using the tools described above.

This gives interesting options for the future. It is important to keep the references between the CIDOC-CRM compliant models and the XML documents they are based on. As eXist is commonly used for storage of XML documents used in cultural heritage through its popularity in the TEI community, see e.g. [PGWR*05], integration between TEI and CIDOC-CRM, which has been researched for some years [EO07] will be easier to implement.

eXist comes with a set of XPath/XQuery interfaces to perform queries over the XML collections and retrieve relevant information in a fast and simple way [XQu05]. A complete transformation framework based on XSL Transformation and a pipeline mechanism allows on-the-fly mapping, transformation, serialisation and presentation of XML content in many other formats, including HTML and RDF [eXi].

7.2. MAD: The SAD extension

The second release of the tool was enriched with a set of features devoted to the management of RDF documents encoded using CIDOC-CRM, the ontology we have chosen for modelling our data. RDF-encoded ontological information required a new query framework implementing SPARQL and RQL, query languages specifically designed for retrieving information from semantic data. The SAD (Semantic MAD) extension was created mainly for this purpose [RDF04a].

Since RDF can also be represented in XML, we tried first of all to use XQuery for the location and retrieval of information from semantic documents. However we noticed that XQuery was insufficient for querying RDF graphs, even when encoded in XML [RDF04b]. An RDF graph is actually a set of triples, each consisting of a subject, an object, and a property relationship between them. These triples can come from a variety of sources. For instance, they may come directly from an RDF document or an ontology instance, they may be inferred from other RDF triples or be the RDF expression of data stored in other formats, such as XML or relational databases.

W3C has defined a new query language called SPARQL specifically designed for getting information from RDF graphs. The language consists of three specifications:

1. the *query language* itself, that is the core of the whole language
2. the *query results XML format*, a simple format easy to process with common XML tools such as XSLT
3. the *data access protocol* used to define simple HTTP and SOAP requests for remotely querying RDF databases and any data repository that can be mapped to the RDF model

The XML results format is used to generate responses from services that implement this protocol [SPA06]. SPARQL features include facilities to:

- extract information in the form of URIs, nodes and literals
- extract RDF subgraphs
- construct new RDF graphs based on information in the queried graphs

As a data access language, it can easily be used for both local and remote access combining the ability of database queries to pull data from huge databases with the possibility to write queries in an application that can extract relevant data stored across the entire World Wide Web. The SAD extension implemented most of these features using APIs provided by the SESAME framework to build semantic query capabilities on top of the native XML database used for storing data.

Another important feature included in the SAD extension is the semantic browser that allows users to navigate through the complexity of semantic relationships.

The richness of the RDF graph model in which semantic data are distributed will often make it difficult for users to get an effective and meaningful data retrieval, if only a simple or complex query interface is used, in particular when the graph grows in dimensions and complexity. Sometimes it would be simpler and faster to browse the multidimensional structure of the graph, allowing users to chose a starting point and to move along different paths through the graph to reach the desired data.

To allow this kind of data navigation we implemented a semantic browser based on the *faceted browsing* user interface paradigm which, unlike a simple hierarchical scheme, gives users the ability to find items based on more than one dimension. A facet is a particular metadata field that is considered important for the dataset that users are browsing. The tool can be configured to prioritise which facets are "open" when the page loads and in what order. Once the facets are selected for a specific dataset, the browser starts processing the dataset and extracts a list of facets, their values, and the number of times each facet value occurs in the dataset. Then it is possible to add or remove restrictions in order to focus on more specific or more general parts of the model.

A "free text" restriction that reduces the browsed dataset to all items that contain the searched string in their properties' values is also possible by entering a search string in an input box.

7.3. Geographic information in MAD

During the last year of EPOCH's activity, we created some experimental geographic functions to integrate spatial archaeological information for the management of unstructured documents, such as excavation diaries and reports, in a spatial context. The system, once fully implemented, will allow the creation and distribution of rich geospatial relationships across the Web and the use of geographic data in a Semantic Web scenario. Tests have been carried out using the Geographic Markup Language (GML) to encode geographic data related to archaeological records and to store them in our container. Data serialised by the MAD system can be directly transformed in SVG or visualised using map server web applications. The flexibility of GML features will also allow the implementation of complex query-on-map functions to visually query and generate dynamic maps. MAD can also host and serialise KML archaeological files to be used in Google Earth and Google Maps applications.

The data integration of both geographic and non geographic data stored in a single container and managed using the same interfaces and the native web compliance provided by MAD will make archaeological information ready to be queried, updated and exchanged over the web, promoting the semantic evolution of geospatial web services.

The development of MAD for the management of both spatial and non spatial archaeological data is indeed the first step towards

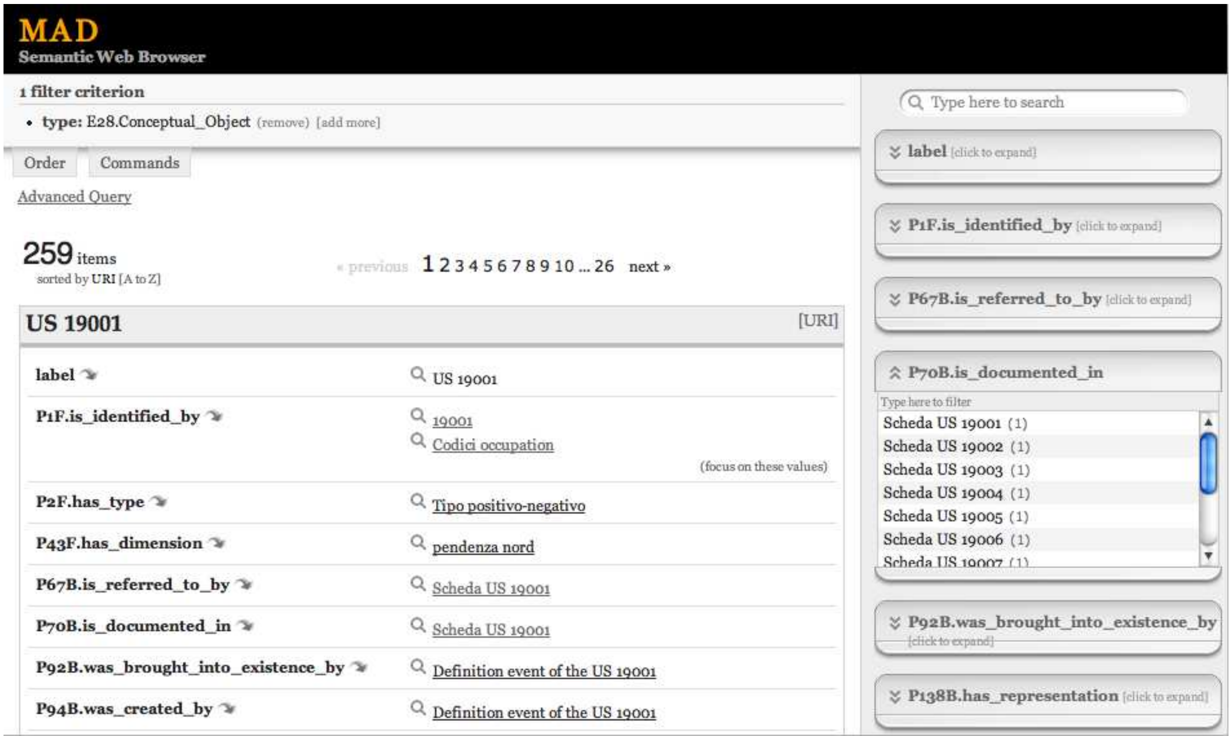


Figure 3: *A snapshot of the semantic web browser of MAD.*

the full implementation of the geospatial semantic web of culture heritage information, a future where the World Wide Web will be machine-readable and fully integrated, allowing the return of both spatial and non-spatial resources to semantic queries.

Integration of spatial and non-spatial information can be also created in MAD during the query process. The power of the MAD query system can return complex sets of XML fragments by recursively executing chains of queries for the generation of aggregated data as shown in the following fragment:

```
<crm:E53.Place rdf:about="US1020">
 <crm:P67B.is_referred_to_by>
  <crm:E73.Information_Object
       rdf:about="gmlModel_US1020">
   <gml:Polygon srsName="osgb:BNG">
    <gml:outerBoundaryIs>
     <gml:LinearRing>
      <gml:coordinates>
       278534.100,187424.700
       278529.250,187430.900
       278528.700,187431.650
       278527.250,187433.600
      </gml:coordinates>
     </gml:LinearRing>
    </gml:outerBoundaryIs>
   </gml:Polygon>
  </crm:E73.Information_Object>
 </crm:E53.Place>
</rdf:RDF>
```

This shows the possibility to create RDFised GML code embedding fragments of GML inside RDF. This capability allows RDF documents to use GML statements for the description of spatial features to be visualised by semantic-enabled browsers with geographic extensions, like the W3C Tabulator [Doe01]. This is in line with the way the TEI guidelines prescribe the inclusion of GML information [TEI08, section 13.3.4.1].

7.4. Mapping support in MAD

MAD was conceived from its very beginning as a repository with a query and a transformation engine on top to provide users with a generic semantic server for their applications. Users simply store their documents, ontologies and models into the MAD database and all this information will immediately become available to be browsed, queried, transformed and exchanged. But the encoding process for data to be stored in MAD is often a complex task for users to accomplish, particularly for those who have to deal with legacy data stored in diverse archives. Even the mapping files created by AMA need further processing operations to be used for data conversion. For this reasons we are going to provide MAD with a mapping framework able to connect the MAD engine with legacy databases and to extract information stored in such databases in an automatic way according to the high level mapping files created using the AMA Mapping Tool. For this purpose we are going to write a new set of Java features based on D2R Server, a tool for publishing relational databases on the Semantic Web working with, and D2R MAP, a simple rather powerful database-to-RDF mapping language [D2R]. D2R Server works with Oracle, MySQL, PostgreSQL and any SQL-92 compatible database. Microsoft Access is partially supported. D2R MAP files (containing specific mapping information) can be easily generated by MAD from the high level mapping files created with AMA.

8. Results and further development

We have tested the MAD framework to build an on-line version of the archaeological dataset recorded during the excavation of the ancient city of Cuma, containing information on stratigraphical units and other related resources and for the creation of an on-line application for the complete management of coins collections for the COINS Project [COI].

The XML native support provided by MAD can be used in the field of digital preservation for setting up annotation repositories and creating co-reference resolution services. It can also assist in the process of linking database type information together with text type information.

In the future, it will be necessary for memory institutions, such as museums, to integrate their data with data from other institutions, crossing thematic, administrative, historical as well as language and cultural borders. To be able to do so, mapping is vital. Such mapping will need clear, educated human minds. But by continuing the development of tools to assist them, we will be able to make the most out of their effort.

References

[AMA] Ama, archive mapper for archaeology. URL: http://www.epoch.eu/AMA/.

[CDG*05] CROFTS N., DOERR M., GILL T., STEAD S., M. S. (Eds.): *Definition of the CIDOC Conceptual Reference Model*. 2005. URL: http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.doc.

[COI] The coins project. URL: http://www.coins-project.eu.

[D2R] D2r server. publishing relational databases on the semantic web. URL: http://www4.wiwiss.fu-berlin.de/bizer/d2r-server/.

[D'A06a] D'ANDREA A.: Documentazione archeologica, standard e trattamento informatico. metodi informatici per la ricerca archeologica. In *Archaeolingua* (Budapest, 2006).

[D'A06b] D'ANDREA A.: Ipotesi di ontologia applicata ai modelli descrittivi/interpretativi dello scavo archeologico. In *Le Ontologie in Campo Umanistico. Archeologia, Architettura e Beni Culturali* (Firenze, 2006).

[D'A06c] D'ANDREA A.: A preliminary ontology-based model applied to the description/interpretation of the archaeological excavation. In *Proceedings of the First International Workshop on "Ontology Based Modelling in The Humanities"* (2006), von Hahn W., Vertan C., (Eds.), University of Hamburg, pp. 38–46.

[DMZ06] D'ANDREA A., MARCHESE G., ZOPPI T.: Ontological modelling for archaeological data. In *The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST2006* (2006), Ioannides M., Arnold D., Niccolucci F., F. M., (Eds.), pp. 211–218.

[Doe01] DOERR M.: *A comparison of the OpenGIS Abstract Specification with the CIDOC CRM 3.2*. Tech. rep., FORTH, 2001.

[EH06] EIDE O., HOLMEN J.: Reading gray literature as texts. semantic mark-up of museum acquisition catalogues, 2006. Paper presented at CIDOC 2006, Gothenburg. URL: http://www.edd.uio.no/artiklar/teknikk_informatikk/CIDOC2006/EIDE_HOLMEN_Reading_Gray_Literature.pdf.

[EO07] EIDE O., ORE C.-E. S.: From tei to a cidoc-crm conforming model. towards a better integration between text collections and other sources of cultural historical documentation, 2007. Poster presented at the Digital Humanities 2007 conference. URL: http://www.edd.uio.no/artiklar/tekstkoding.html.

[eXi] exist open source native xml database. URL: http://exist-db.org/.

[FRB98] *Functional Requirement for Bibliographic Records. Final Report*. Tech. rep., IFLA, 1998. URL: http://www.ifla.org/VII/s13/frbr/frbr.pdf.

[HOE04] HOLMEN J., ORE C.-E., EIDE O.: Documenting two histories at once: Digging into archaeology. In *Enter the Past. The E-way into the Four Dimensions of Cultural Heritage* (2004), BAR International Series 1227, BAR, pp. 221–224.

[JHOOng] JORDAL E., HOLMEN J., OLSEN S. A., ORE C.-E.: From xml-tagged acquisition catalogues to an event-based relational database. In *Proceedings from CAA 2004* (forthcoming), BAR.

[PGWR*05] PETTER C., GROVE-WHITE E., ROBERTS L., ROSE S., POSGATE J., SHOICHET J.: The robert graves diary (1935-39): a tei application using an xml database (exist). In *ACH / ALLC 2005. Conference Abstracts (2nd Edition)* (2005), University of Victoria. Humanities Computing and Media Centre, pp. 164–166.

[RDF04a] Resource description framework (rdf): Concepts and abstract syntax, recommendation, 2004. URL: http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/.

[RDF04b] Rdf/xml syntax specification (revised), recommendation, 2004. URL: http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/.

[SPA06] Sparql query language for rdf, candidate recommendation, 2006. URL: http://www.w3.org/TR/2006/CR-rdf-sparql-query-20060406/.

[TD01] THEODORIDOU M., DOERR M.: *Mapping of the Encoded Archival Description DTD Element Set to the CIDOC CRM*. Tech. rep., FORTH, 2001. URL: http://cidoc.ics.forth.gr/docs/ead.pdf.

[TEI08] *TEI P5: Guidelines for Electronic Text Encoding and Interchange. Version 1.0.1. Last updted on 3rd. February 2008*. TEI Consortium, 2008. http://www.tei-c.org/Guidelines/P5/.

[XQu05] Xquery 1.0: An xml query language, candidate recommendation, 2005. URL: http://www.w3.org/TR/2005/CR-xquery-20051103/.

Interpretation Management: How to make sustainable visualisations of the past

Daniel Pletinckx

Visual Dimension bvba, Ename, Belgium

Abstract

Current technology allows us to more and more easily create threedimensional models of landscapes and man-made structures and to visualise these models in several interactive and non-interactive ways. In the eighties, the idea arose at IBM to use this technology, which had been developed for designing and visualising structures that still had to be built, also for visualisation of structures that had existed but disappeared for one reason or another.

Although there is no fundamental technological difference between visualising structures that still need to be built and structures that have existed, there is a major conceptual difference because our knowledge of the past is partial and uncertain. In fact, we are not able to reconstruct the past at all. Even for the nearby past, we lack a lot of information to fully reconstruct structures that have disappeared.

We can try to puzzle together all information we have about a certain structure in a certain time period, and try to visualise this incomplete and uncertain information in the best possible way. This paper explains the methodology for doing this in a correct and reproducable way. In fact, archaeological and historical research have been using similar methods already for a long time, but this methodology hasn't been implemented yet for 3D visualisation, except for some pioneering efforts (see for example [NUME], [ROME]).

In this paper, we explain and illustrate methods such as source assessment, source correlation and hypothesis trees that help to structure and document the transformation process from source material to 3D visualisation. We will also discuss the different approaches of 3D visualisation in research and in public presentations, and present a tool to manage the interpretation process.

The key goal of this paper is to propose a methodology and tool to make open, sustainable 3D visualisations of the past and turn them into an instrument that is accepted in both the research and public presentation domain. This tool is part of the EPOCH Common Infrastructure that provides concrete tools and solutions for common problems in the cultural heritage domain.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Methodologies and Techniques]: Standards

1. Background

3D visualisation uses our current capabilities to create three-dimensional models of objects, and show them in different ways with varying degrees of realism and interactivity. 3D visualisation has proven to be able to recreate and visualise historical structures (buildings, cities, landscapes, man-made structures, ...) and is becoming more and more an accepted method for showing interpretation in historical and archaeological research.

3D visualisation however had and partially still has the connotation of lacking crediblity and documentation [RYAN96], which can lead to producing too optimistic and even false conclusions about the past and about the premises and possibilities of archaeology as a discipline [RYAN01]. Many cultural heritage specialists have voiced their concerns about the improper use of 3D visualisation and the lack of a proper methodology to produce recreations of the past [BOU00].

Although the term *interpretation* has other meanings and connotations in other domains, we use it here to indicate the intellectual process of turning source material into conclusions. In the context of 3D visualisation, these conclusions are of course focused on all visualisation aspects, but in fact the visualisation process is embedded in most cases in the wider interpretation process, and only helps to support research or transfer knowledge to the wider public.

We don't use the term *virtual reconstruction* because our main goal is not to reconstruct the past – this is something we simply can't do – but to bring together all available sources of information and visualise this with 3D technology. Visualisation can be very useful in a research context but also for public presentation. This means that we don't want to obtain always photorealistic, complete models of landscapes or man-made structures, sometimes we only want schematic or simplified representations. Therefore we use the general term *3D visualisation*.

Most of the technological issues in this field have reached a sufficient level of solution, and a variety of tools is available for most 3D visualisation tasks. The process of turning available sources into a 3D visualisation on the other hand is far less defined. This interpretation process not only takes most of the time within the visualisation process, it is also a complex, non-linear process that can profit significantly from tools that manage and organise this process. In other words, *interpretation management* is a key element of 3D visualisation of historical structures, as it records and manages how the available sources have led to the 3D visualisation, and supports and smoothes the interpretation process.

2. What purpose does interpretation management serve?

There are several reasons why interpretation management is necessary when visualising 3D models of historical structures.

First of all, it *records* the interpretation process and documents how all elements in the visualisation have been derived from the available sources. This is a necessary step, as practice shows that 80 to 90 percent of the work of 3D visualisation of historical structures goes into the assessment and interpretation of the sources, only 10 to 20 percent of the time is spent on building the 3D model. Practice learns that this interpretation process is complex and can extend over a long period, that the amount of source data can be overwhelmingly large, and that in many cases multiple people work simultaneous on the same project. Following well defined procedures, supported by a tool that records and manages this interpretation process, is therefore crucial in safeguarding the majority of the financial and intellectual investment of a visualisation effort.

A second reason for having interpretation management is the ability to *update* 3D visualisations with new results, coming from new excavations or recently discovered historical sources or from new scientific interpretations and insights. The influence of such new data is in most cases far from straight forward, hence having a well defined process how new results alter the interpretation process is necessary to properly manage existing 3D visualisations. In other words, 3D visualisations should remain “alive”, even many years after excavations or research efforts have ended.

This brings us to a third element which is *scholarly transparency*. When visualising historical buildings or landscapes, we need a lot of information to build complete 3D models. In most cases, we have insufficient and indirect sources to construct the 3D model, so coming from those available sources to a complete 3D model is a difficult process. We have to understand that the uncertainty of elements in a 3D visualisation can vary largely across the model, some elements are well defined while some elements are totally unclear. The process of how to fill in these uncertainties is undefined, and can yield several good solutions. Even more, when basic choices are unclear (e.g. is the excavated structure a small church or a large house ?), results can depend to a large extent on small details or even speculations or assumptions. This means that many 3D visualisations, or at least parts of it, can have large amount of uncertainty. For public presentations, it is not always useful to expose this uncertainty, hence a certain choice on what and how to show will be made, but for scientific purposes, a 3D visualisation needs to be transparent, and the uncertainty and choices made need to be well documented, and available for scientific critique and research. In other words, interpretation management is a way to “publish” 3D visualisation.

A fourth element is *data security*. Practice shows that most visualisation processes yield binders of unstructured documents from which outsiders cannot reconstruct the interpretation process. In other words, the intellectual efforts linked to creating a 3D visualisation cannot be passed onto the next generations. By providing a methodology and tool to record and manage the interpretation process of a 3D visualisation in a structured way, we also provide a way to store this data for the long term, giving access to the data and the interpretation process for future use and research.

A final element is *multidisciplinary cooperation*. We need to keep in mind that 3D visualisation brings together a wide range of skills (from history and archaeology to architecture and stability engineering, from pollen analysis and hydrography to 3D modelling and rendering) and that it is impossible that one person can master all the skills needed to do proper interpretation of all available sources. A tool that brings together all sources and all interpretations is in fact also a *collaboration platform* that allows all involved disciplines to contribute their part to the project, mainly in an iterative process.

3. Recording methodology

This tool wants to be practical and usable and helps supporting the 3D visualisation process. We need to be aware that this tool needs to be simple, create nearly no overhead and needs to adapt itself to a large range of situations.

We need to be aware that not many such tools have been introduced to the 3D visualisation community yet and that only practical use by a large number of experienced people will show how the tool needs to be further developed. Hence, we think that it is wrong to be too prescriptive and too restrictive by forcing people into a rigorous framework. The tool should rather be a container where information can be stored in a flexible way, gently guiding people through the interpretation process following the lines of a certain methodology.

The methodology for interpretation management presented here is based upon many years of experience in 3D visualisation. The main features of the methodology are:

- clear references to all sources used, no use of implicit knowledge
- in-depth source assessment, making the reliability and potential bias of each source clear
- correlation of all sources used for a certain visualisation in order to detect common ground as well as inconsistencies, outliers or dependencies
- structural analysis of the object to be visualised, and division of the object into logical sub-units
- list of all potential hypotheses, never “hiding” a discarded hypothesis

- records the interpretation process by making a clear link between the sources, the reasoning and the resulting hypothesis
- structures the potential hypotheses in a tree structure, with sub-hypotheses depending on main hypotheses
- keeps the recording process separate from the modelling and visualisation process, as the latter is far from linear

It's the rigorous implementation of this methodology in general and the use of correlation techniques for iconographic sources and a hypothesis tree in particular that makes it well suited to optimise the process of constructing a virtual model from related sources.

The methodology we propose here is basically a step-by-step process:

1. Creating a source database
2. Source assessment
3. Source correlation
4. Creating hypothesis trees with conclusions
5. Updating

We deal with issues such as the reliability of the hypotheses, multiple hypotheses with the same level of probability, ways to express uncertainties and visualising evolution. We explain the goal and approach of the London Charter, and demonstrate how the EPOCH tool implements these goals.

3.1. Creating a source database

It is a good practice to refer systematically to sources, and document these sources through references, images and text descriptions (many people forget that text is one of the most important sources in a 3D visualisation process). These sources are maintained in a *source database*. Practice shows that many errors in 3D visualisation are due to incorrect assumptions when using source material. Having a rigorous process to select and document sources helps avoiding this pitfall.

There is no standard way to structure the source database, as many different types of sources can be integrated (from iconography to pollen analysis, from unpublished excavation data to well-known historical sources, from historical analysis of existing buildings to oral history). The principle needs to be that *all sources are identified uniquely and can be traced easily when needed*. Basically, this does not differ from standard practice in archaeological and historical research (where footnotes are used in most cases) but more technically oriented people making 3D models need to adopt this too.

Each source is referenced on a source sheet that also contains digital images, details of those images or transcriptions of text where necessary. Having such key information copied in the interpretation management system is very useful to avoid physiscal search in documents, which can be available in libraries and archives only.

3.2. Source assessment

A key element in the interpretation process is *source assessment*. This assessment normally yields some understanding of the reliability of the source, and more specifically about the reasons why certain elements are not reliable.

This *assessment* can be a detailed study of the context of the source or the way the source depicts the reality. For example, iconography needs to be studied in terms of the creator of the iconography, the reason why the iconography was made or how the iconography needs to be interpreted. In fact, source assessment tries to know and understand the process how reality was represented in the source at hand.

We need also to be aware that all sources, from text sources or iconography to archaeological sources or digitised buildings and objects, have been interpreted already during their creation, hence that mistakes, missing information, incorrect interpretations or deliberate alterations can occur, and that we need to understand the context of the creation of the source to try to get the maximum of correct information out of the source. By applying correlation with other independent sources (see next step) we can try to further remove the veil of error that is present in every source.

3.3. Source correlation

The *correlation method* compares the different sources and tries to draw conclusions from the corrspondences, differences and inconsistencies between the sources. Conclusions can be that a source is totally unreliable, contains certain deliberate errors or just mistakes, or is a correct and detailed representation of the item it depicts or describes.

The basic correlation method is *consistency checking* between sources that basically contain the same information. This can for example happen between different sources of iconography depicting the same scene, or archaeological sources versus iconography. Of course, it is important to keep the context in mind as a drawing from the middle ages for example cannot be expected to contain proper perspective. We also need to take the character and limitations of the sources (as recorded in the source assessment) into account.

A special case of this consistency checking is when *several versions* of a certain source exist. By analysing small differences between the different versions, and by historical study, in most cases the most reliable (often the oldest) source can be identified.

In most cases, we don't have the luck to find multiple sources such as drawings or paintings that basically depict the same. Normally we have different types of sources that depict the same environment at different points in time, made for different purposes. Correlation in that case consists of a *systematical comparison* of all available elements, record common elements and try to understand why some elements are different or absent. As the major hurdle to take here is understanding the evolution of the structure, we need to correlate all available sources on that structure at once (see chapter "Visualising evolution" below).

We have analysed several methodologies to formalise this correlation process, but as this is a very non-linear and complex process, finally it seems that only *description through text* can capture all the necessary nuances and be adopted easily. The short description of the tool below gives a good idea how this is done.

3.4. Making a hypothesis tree with conclusions

When visualising a building, a landscape or a city, we need to impose a certain top-down analysis of the object, decomposing it in *substructures*. These substructures do not always follow the normal, "structural" decomposition of the object but rather the logical decomposition, hence they are closely linked with the hypothesis tree we will introduce. Nevertheless, the object needs to remain well structured and plausible. Creating too much structure where no information is available generates only an additional burden for the person making the visualisation, we need to keep in mind that the methodology needs to support the visualisation process, not making it more complex.

The hypothesis tree is the *formalisation of the interpretation process*. It shows in a top-down fashion the potential alternatives, analyses each of the alternatives in relation to the available sources and draws a *conclusion* about which one of the alternatives has the highest probability, based upon the available sources.

In each hypothesis, *sub-hypotheses* are made, which again are evaluated and the most probable one is selected. The reasoning how the sources (indicated through hyperlinks) influence the hypothesis is done in written text, we do not believe a formal structure can be devised that is both flexible and user friendly enough to refrain from the normal written word, that everybody uses to express interpretation.

It is important though to stick to the branching hypothesis tree method, to avoid overlooking certain possibilities. Nevertheless, it is common sense that unlikely branches do not need to be expanded as this only creates additional overhead that is not useful, but the unlikely branch needs to be recorded anyway (see updating methodology).

A hypothesis tree implies at first sight that the analysis happens in a top-down fashion. For this, all information needs to be available, so related excavations and historical studies have to finalised. Archaeologists on the other hand want to work in a bottom-up fashion while they excavate and can only merge parts of structures to complete structures when excavations finish. Hence, the tool we will use to document the interpretation needs to be able to deal with this workflow in an elegant way.

Most historical structures show an *evolution through time*. When interpreting source data and proposing certain hypotheses, we need to think in fact in four dimensions, spatially and chronologically. In other words, every hypothesis needs also to check if it is consistent with the data of the phases before and after a specific 3D visualisation. Arriving at a consistent evolution is a major part of the interpretation to be done, and a major validation step when building or updating the virtual models.

Therefore it is important to entangle the different *phases* of a structure, in other words, interpretations should cover the full evolution of a building, landscape or site. Of course, when there is a discontinuous evolution (for example, a site is demolished and rebuilt in a totally different way), the interpretation can be divided in those discontinuous phases, and be treated separately.

3.5. Updating

One of the most important reasons to do interpretation management is *updating*. As new sources of information can appear, as new insights or correlations can be found during the study of the source material, we need to be able to record how this new material influences the existing 3D visualisations. We distinguish four different kinds of updating.

First of all, when a *new source* appears, we need to add this source to the database, find out what other sources it correlates to and assess this new source, both on its own and in comparison to all other related sources. The availability of new source material can influence the assessment of other sources, the reliability of the visualisations or even the hypotheses made (see below).

Another update action is the appearance of a *new assessment of an existing source* where new insights, new sources or new studies (which need to be added to the source list) render the current assessment of a source obsolete or at least incomplete. This new assessment can trigger changes in the hypotheses section and of the reliability of the visualisations.

New sources, changes in source assessment or new interpretations can yield an *additional or updated hypothesis* or can *change the probability of one or more hypotheses* or *the reliability of the visualisations*. This can yield in another conclusion (the hypothesis that has the highest probability) than before.

In this process of updating, there needs to be a detailed *tracking* of the updates. This is not only a technical issue, there needs to be a *consensus* amongst the involved people on any changes to the 3D visualisation, and the changes need to be implemented and validated by 3D specialists. As pointed out before, this normally is an iterative process that needs involvement of several specialists, leading to a change to the virtual model by the 3D specialist. As in most cases these specialists do not share the same working space or meet each other daily, we need a tool that can act as an internet collaboration platform to allow these interactions to take place efficiently.

It can happen that specialists do no agree on a certain conclusion, or that too little evidence is present to favour one interpretation over another, or that the update is not endorsed by all involved specialists. In that case, there are *two or more solutions* that are treated *as equally probable*. This is in itself not problematic, but needs in-depth consultation and consideration before the decision can be taken that there is no most probable interpretation and 3D visualisation.

It is clear that a certain degree of skills is needed to make or change the interpretation and visualisation of a site. This is the same problem as Wikipedia is facing to maintain the quality of its online encyclopaedia and avoid "vandalism" of the content. Like Wikipedia, everybody needs to be able to contribute to the interpretation of the sources, following the typical discussion methodology and user authentication. Unlike Wikipedia, there should be an authorisation and accreditation process of people who want to change the conclusions and make or change the 3D visualisations, as these are complex tasks that require the appropriate skills. These accredited specialists can be seen as the "scientific committee" of the 3D visualisation programme. We think we can guarantee in this way the quality of a 3D visualisation while "publishing" this visualisation and creating full transparency about the interpretation.

All data that is stored as result of the creation and update process also needs a maintenance cycle that should not be longer than two years. The software of the implementation (see below) and its associated data (typically a database with all results) probably will need to be updated. Files integrated in the database (such as digital images) or in a digital repository (3D virtual models, derived results such as animations, interactive models, ...) need to be transferred to new file formats if the original file formats become obsolete (this is called "data migration").

3.6. The reliability of the hypotheses

Besides what is most probable, we also need to care about the reliability of the visualisations that result from the most probable hypotheses. Although it is difficult to put a number on the reliability of each structural element of a visualisation, we can derive some estimation from the reliability of the sources (see source assessment and source correlation) and the number of sources that are available for that specific element (see source correlation). In most cases, an indication of high, medium and low reliability is sufficient. If we have only unreliable sources or if we only have one source, we will attribute the visualisation a low reliability. If we have multiple, reliable sources, we will consider the visualisation as highly reliable.

In the same way, if a hypothesis matches perfectly with all available sources, the visualisation can be considered as highly reliable, while if a hypothesis matches poorly with the available sources, but no better hypothesis can be found for the moment, the visualisation needs to be considered as unreliable (even if the hypothesis is considered most probable).

Unlike some other specialists in the field of 3D visualisation [HER05], we prefer not to quantify reliability in numbers but assess the reliability as *low, medium or high*. Other authors use a similar methodology. Peter Sterckx [STE07] uses the same system for the visualisation of the evolution of the Horst castle in Belgium, while Han Vandevyvere [VAN06] uses four catergories (low, medium, high and very high) for the Mariemont castle, as does Matt Jones [JON07] in his visualisation of Southampton in 1454.

The issue however is what to do with unreliable parts of the visualisation. Should we visualise them or not? When we start from a scholarly point of view, we rather will decide to not visualise unreliable parts. When we start from a presentation point of view, we try to show a *consistent* image of the visualised structure, so we rather will decide to show also the unreliable parts because they make the structure as a whole more consistent.

3.7. Dealing with multiple hypotheses with the same level of probability

If one hypothesis has clearly a higher probability than the others, the *conclusion* will put this hypothesis forward as the most probable interpretation of the available sources. However, if two or more hypotheses have more or less equal probabilities, the conclusion needs to reflect the undecided nature of the interpretation. In that case, all probable alternatives will be expanded, i.e. will have sub-hypotheses and developed virtual models.

Nevertheless, if the alternatives are not significantly different, one hypothesis can be chosen as the *representative conclusion* for public presentation, regarded that information is available in that presentation about the other equally probable alternatives.

3.8. Visualising evolution

When visualising evolution, we basically want to explore a 3D structure from all sides and see the evolution of (a part of) that structure from the most appropriate angle.

Several technical solutions have the potential to do that, but we want to present here a simple but very powerful techique: a QuickTime VR object. QuickTime VR [QTVR] is part of the QuickTime software that is able to visualise panoramical and spherical images and interactive objects.

Interactive objects basically consist of a matrix of images that can be visualised interactively by dragging horizontally or vertically in the viewer. If we put a 360-degree rotation of the object in the horizontal rows of the matrix, and an evolution through time in the vertical columns of the matrix, then we obtain a *4D visualisation tool* that shows interactively 3D plus time (evolution). Hence, if we drag our cursor horizontally or use the left/right arrow keys, we change our viewpoint, while if we drag vertically or use the up/down arrow keys, we visualise the evolution of the object from a particular point of view.

Simple software packages exist to turn a set of images, structured in such a matrix like 4D way, into such an interactive 4D object. The major advantage is that from the interactive object, *hyperlinks* can be made so that it can be integrated into hyperlink-based tools.

3.9. The London Charter

The London Charter [TLC] has been initiated at a meeting of 3D visualisation specialists in 2006 in London and aims to define the basic objectives and principles of the use of 3D visualisation methods in relation to intellectual integrity, reliability, transparency, documentation, standards, sustainability and access. It recognises that the range of available 3D visualisation methods is constantly increasing, and that these methods can be applied to address an equally expanding range of research aims.

The Charter therefore does not seek to prescribe specific aims or methods, but rather seeks to establish those broad principles for the use, in research and communication of cultural heritage, of 3D visualisation upon which the intellectual integrity of such methods and outcomes depend.

The Charter does seek to enhance the rigour with which 3D visualisation methods and outcomes are used and evaluated in the research and communication of cultural heritage, thereby promoting understanding of such methods and outcomes and enabling them to contribute more fully and authoritatively to this domain.

So the London Charter can be seen as the upcoming standard for 3D visualisation. The methodology we propose here is a way to implement the Charter (version 1.1) in practice, which is based on the following principles [TLC]:

- valid for 3D visualisation in all cultural heritage domains
- appropriate use of 3D visualisation
- identification and evaluation of relevant sources
- transparency of the 3D outcomes in relation to the sources
- documentation of the 3D visualisation process should allow repeatability of the interpretation process and reuse of the outcomes, and create a scientific dialogue and understanding
- use of standards and ontologies, approved by the community
- sustainability
- improve accessibility of cultural heritage

The London Charter wants to be valid for *all domains* in which 3D visualisation can be applied to cultural heritage. This tool is also very general and has a methodology that can be applied in a wide range of applications and for a wide range of goals. The concept of assessing sources before they are used in the interpretation process, and the method of correlating sources to reveal common thruth are generally applicable. The concept of building a tree of hypotheses allows to work both bottom-up and top-down, which makes it suitable for a wide range of cases. The methodology presented here can be used for research as well as for communication purposes.

The London Charter states that an evaluation of the goals to achieve should prove first of all if 3D visualisation is an *appropriate method*, and if so, which 3D visualisation method is the most adequate to reach the goals. The methodology used here is quite independent of the visualisation method of the results of the interpretation process, which could range from 2D maps over pencil sketches to 3D volume renderings and photorealistic 3D visualisation.

The London Charter states that *sources should be identified and evaluated in a structured way*. This is exactly one of the key elements of this methodology, and is explained and demonstrated in detail in this paper.

The London Charter states that the relation between the sources and the 3D visualisation outcomes, the reliability of those outcomes and the interpretation process should be *transparent* and well documented. This again is one of the key elements of this methodology, and is explained and demonstrated in detail in this paper.

The London Charter promotes the *scientific rigour of the interpretation process*, based upon documentation, reuse of results and scientific dialogue. These elements are also key elements of this methodology and its implementation through wiki-technology.

The London Charter promotes the development of *standards and ontologies* for documenting the interpretation and 3D visualisation process and enventual approval by the community that develops and uses these 3D visualisations. The proposed methodology in this text and its implementation through a wiki-based tool, together with possibly other implementations, can be considered as a first step and platform for the community to develop these standards and ontologies, as this is still a new and uncharted domain.

The London Charter promotes *long-term archival and sustainability* of the documentation of the interpretation process and its resulting 3D visualisations. This is also a major goal of the methodology proposed in this paper. The use of a standard wiki-tool and its inherent mechanisms of archival and versioning will help in realising this goal.

The London Charter states that the documentation of the interpretation process and 3D visualisation outcomes should provide a better *access* to cultural heritage assets in terms of study, interpretation and management. The approach of a wiki-based tool that is accessible for authorised specialists allows a high quality enviroment that can be used for further study, scientific discussion about interpretation and documentation for a wide range of uses.

5. Structure of the tool

EPOCH, as the Network of Excellence for the use of ICT in cultural heritage, has created tools for the cultural heritage community to support specific tasks [EPOCH]. For 3D visualisation, a tool based on the methodology explained in this paper, has been created and is freely available.

The tool has five major functionalities: the source database, source assessment, source correlation, the hypotheses tree with conclusions and the 4D visualisation page. It is based upon wiki technology that implements not only the hyperlinking, but also the discussion forum and the consensus process that is needed to communicate and discuss research results and update them when necessary. Resulting 3D models or derived products (still images, animations, ...) can be stored in a data repository and hyperlinked to the 4D visualisation page.

5.1. The source sheet

The *source database* needs to refer to the sources used in the interpretation. Strictly speaking, these sources need only to be traceable. In practice, storing enough data in a digital way makes the sources much more accessible, hence supports the interpretation process. For example, if iconographic sources are available as high-resolution images, this is of course more practical when doing interpretation than juggling with a lot of printouts and photographs. In other words, appropriate viewers, tools to find the right sources and good management of multiple windows are important. A functionality that is very powerful is the use of *annotated images*. In this way, text such as transscriptions or remarks can be added to certain areas of the images and appear when the cursor moves over the annotated area.

In this way, each source is recorded on a *source sheet* that also contains the *source assessment* that records the context, quality and interpretation of the source in itself.

We distinguish *primary and secondary sources*, the former are directly used in the interpretation process, the latter are only used in the source validation and don't need to have images, validation, ... (if there would be an issue with reliability of these secondary sources, a source assessment section can be added).

5.2. The source correlation sheet

On the *source correlation sheet*, we document the correlation process between sources that give us information about a certain topic. By numbering features in iconographic sources that match, we can extract elements with a higher reliability that can be used in the visualisation process.

A more elaborate version of this, that is used when there are significant differences between sources, is the technique where matching features are indicated in green while features that do not match are indicated in red. Features that have an uncertain matching can be indicated in yellow. Based on such an analysis, a detailed description will be recorded in the source correlation sheet that focuses not only on the matching features but also on the differences, and why they are different.

In this process, the secondary sources, that are referenced to provide the context of the source, simply go into the source database. The source correlation is captured in text with hyperlinks referring to the sources and annotations to the images recording remarks, transcriptions or interpretations.

5.3. The hypothesis sheet

The *hypothesis tree* provides a simple way to decompose the interpretation of a structure (building, landscape, ...) into a tree-like set of possibilities called hypotheses. Each hypothesis can be decomposed in other sub-hypotheses. At each level of the hypothesis tree, conclusions are made, based upon the probability of the possible hypotheses. If a hypothesis has little probability, one can choose not to generate sub-hypotheses for it.

We propose here to build this hypothesis tree as a set of hyperlinked pages called *hypothesis sheets*. This approach gives the freedom to change the structure of the tree easily. One common problem in historical interpretation is the assumption of certain conditions or elements. If we want to remove the assumption that is made at a certain level, we need to add an extra level of subhypotheses in the middle of the tree, which is very easy by changing a few hyperlinks.

As archaeologists like to work bottom-up instead of top-down (as they are dependent on what has been excavated already), they can make hypothesis sheets that deal with parts of structures and link them later on with overarching hypotheses once the excavations have progressed or finished. The flexibility of creating and modifying the hypothesis tree through creating or modifying hyperlinks makes the approach described here applicable in many different cases and contexts.

The structure of the hypothesis tree also reflects the logical structure of the object to visualise (see chapter 3.4). As Vatanen [VAT03] points out, only hyperlinked structures can provide the necessary flexibility and interlinking to represent three-dimensional structures (as they can show a complex evolution pattern, they can even be considered as four-dimensional).

If one hypothesis has clearly a higher probability than the others, the *conclusion* will put this hypothesis forward as the most probable interpretation of the available sources. Each conclusion is linked to the *visualisation of the resulting virtual model* and appropriate tools refer to parts of the virtual model, linked to each sub-hypothesis.

If two or more hypotheses have more or less equal probabilities, the conclusion will reflect the undecided nature of the interpretation. In that case, all probable alternatives will be expanded, i.e. will have sub-hypotheses and developed virtual models.

Nevertheless, one hypothesis can be chosen as the *representative conclusion* for public presentation, regarded that information is available in that presentation about the other equally probable alternatives.

Each conclusion should also make an estimation of the *reliability* of the most probable hypothesis. This depends on the number and quality of the involved sources and on the matching between the hypothesis and the sources.

Although parts of a structure can have a low reliability, we can choose to visualise those parts to maintain the consistency of the visualisation. The decision to visualise or not visualise depends very much on the purpose of the visualisation. When used in scientific research, it could better to not visualise unreliable parts.

5.4. The 3D/4D visualisation page

In most cases, the interpretation of a structure (building, landscape, city, ...) deals with the evolution of that structure through a certain period of time.

A potential user interface to show this evolution properly can be *4D visualisation* through for example QuickTime VR, where both time and 3D viewpoint can be changed. In this way, links can be made towards the pages of the interpretation, so that the 4D visualisation acts as a user interface to reach the appropriate parts of the interpretation. In other words, clicking on the reconstructed element refers to the exact place in the hypothesis tree, where the structure of that element is discussed.

If the visualisation contains no evolution, this page simply shows the reconstructed element from all sides, in 3D.

5.5. Advantages of a wiki implementation

As it is important to be able to work on a reconstruction as a full team, with the appropriate discussion and consensus mechanisms, we propose to use *wiki-technology* with predefined templates, revision control and consensus mechanisms to build the three components outlined above (source database, source correlation and hypothesis tree). 4D visualisation tools that can be embedded into a wiki provide the fourth component. User registration, authentication and authorization on different levels (edit, conclusions, 3D model creation and edit) can use the tools that most wikis provide. The update of 3D models is triggered by a change of the conclusions, therefore a change in conclusion and a change of the virtual models can only be decided by accredited people.

Another aspect that we try to realise is data security. We try not only to provide a safe, long term storage of the interpretation process, but also long term storage of the virtual models themselves. As the interpretation process and its updating is linked to a specific 3D model, this model can also be stored (together with all related files, such as textures and animation scripts, and derived files, such as images, animations, interactive applications). In this way, a *central, public storage* is implemented so that the investment of making the virtual model is secured. As 3D model files can have a variety of file types, it is important to provide long term maintenance of those files by translating them into open file formats and provide a further translation of those file if the chosen open file format becomes obsolete in the future.

As nearly all virtual reconstructions are made with public money, it is conceivable to request that 3D models are published and maintained in such a way. In this way, there is not only a much better safeguarding of the financial and intellectual effort that goes into the virtual reconstruction project, but also a public availability of the results (as they have been made with public money). The Network of Expertise, that has been created in the EPOCH project, and that hopes to continue under 7FP, will provide training to have these concepts widely accepted and do lobbying to turn this approach into national policies.

6. Other approaches

The Architectural History and Conservation Research group and the CAAD research group at the University of Leuven have defined a metafile approach for documenting 3D visualisations of historical buildings [VAN06]. This metafile has the form of a spreadsheet and distinguishes the logical structures in the object to visualise (rows) and contains the facts, the written and iconographical sources, the onsite inspection reports, reliability and remarks (columns). This metafile also decomposes the different major elements into a tree structure through a numbering system of the rows.

Practical application to complex structures such as the Horst castle in Belgium [STE07] shows however that this approach has unsufficient flexibility to represent the large amount of links between the elements and the large amount of description and interpretation needed for each of those elements. There is no entry for hypotheses which results in the fact that the interpretation gets scattered over the facts column (where it definitely does not belong) and the remarks column. Especially the interlinking between the different hypotheses get somewhat lost in the limited area that each facts or remarks cell provides. The assessment of the sources is not systematic as it ends up in the remarks column, separated from the different source columns. Most sources don’t get any source assessment and there is nearly no source correlation. There is only a text link to the sources, which results in a lot of manual browsing and searching.

Most projects where the metafile approach has been applied are standing buildings for which resolution of interpretation issues at the detail level are the most important. The metafile approach works quite well for this kind of projects but is less optimal for other projects where there are only ruins or archaeological remains left, or when significant rebuilding has take place as in that case, significantly different hypotheses need to be compared to each other.

Another approach has been used by Joyce Wittur [WIT08] in studying the methodology for the Lorsch Abbey Reconstruction and Information System [LARIS]. The Lorsch abbey is a major 8th century abbey and UNESCO World Heritage Site nearby Mannheim in Germany, which has a few standing buildings but also major archaeological remains. This approach uses argumentation networks that link sources with interpretations and is very similar to the approach we take in this paper. These networks are built of interlinked primary data, comparisons, observations and interpretations. The major difference however is that there are no phases in the interpretation process such as source assessment, source correlation and hypothesis building, and that the hypothesis building doesn’t need to have a tree structure.

Each of the elements in the argumentation network can be linked to data such as images, texts, etc. This is very similar to the hyperlink approach we use in this paper and produces an efficient way to link and structure the data.

Although there are no formal phases of source assessment, source correlation or hypothesis construction, all these elements are present in these argumentation networks. The network approach allows on one hand the flexibility needed to describe all kinds of different interpretation processes, but lacks on the other hand the rigour and guiding of the step by step approach that we promote here.

At this moment, the argumentation networks do not care about reliability of sources, the probability of hypotheses or the update process, although we think that these elements are crucial in good interpretation management.

This argumentation network approach has also been promoted by Vatanen [VAT03], who does take care about the history of the interpretation process and updating.

Argumentation networks do represent very well the way we are thinking, but look somewhat complex and scary at first sight. This could influence negatively the take-up by the 3D visualisation community, so we promote in a first phase the use of plain text to describe the arguments. As Vatanen [VAT03] points out, the first goal of documenting interpretation should be communication within the involved community. Nevertheless, they are a good basis for structured storage of the interpretation process in the near future.

7. Wiki template

This section provides a detailed proposal of the structure of the different types of wiki pages (sources with their assessment, source correlation, hypothesis tree, 4D visualisation, overview).

The source database is implemented as a set of *source pages*, which provide the right reference to the physical source, useful tools such as images that are helpful in the interpretation process and a validation of the source itself.

The goal of source assessment is to define if the source is has a certain degree of reliability and therefore is eligible to be used in the interpretation process, and to reveal the interpretation processes that have already taken place in the creation of the source, so that we can define which elements of the source can be trusted and which elements should be discarded.

A source page contains:

- a short description of the source
- a conclusion (of the validation/assessment process) at the top of the page
- a correct and complete reference to the source, and a URL if online
- images (plus details if necessary) related to the source, for practical reference and for supporting the reconstruction process
- a detailed description of the source
- a detailed description of the context of the source (who made it, why, ...), this can require links to other sources (both primary and secondary sources)
- a validation and assessment of the source itself concerning its context and reliability (both through primary and secondary sources)

A source correlation page groups a number of sources that have something in common, and records the correlation process between all these sources, yielding conclusions that can be made from that group of sources. In this way, archaeological sources are correlated with iconographic sources, or similar iconographic sources are compared amongst themselves.

A source correlation page contains:

- a conclusion at the top of the page
- a list of related sources that document a certain structure, area or period, each source is identified through a hyperlink to its own source page (a source can appear in more than one list)
- a comparison of the listed sources, using correlation techniques and based upon the source assessment, trying to assess similarities (are they really similar?), differences (why is there a difference?), omissions (why is a certain structure missing in a certain source?) or interpretations from other authors in relation to the full set of available sources, each source is hyperlinked for easy reference
- drawing conclusions from the correlation process: if all sources depict the same feature in the same way, the conclusion is that the sources show a reliable feature, if the sources show different and non-consistent depictions of a certain feature, we can conclude that we have little to no evidence for that feature

The *hypothesis tree* is implemented by a set of hyperlinked hypothesis pages and each page covers one branch of the hypothesis tree. All hypothesis pages are linked with each other through hyperlinks to form a tree.

A hypothesis page contains:

- one or more conclusions at the top of the page, linked to the 3D/4D visualisation page if available
- all hypotheses concerning a certain structure or part of a structure, each hypothesis links to its sub-hypotheses
- a written description of the argumentation, pros and cons of a certain hypothesis, linked to the appropriate source and source correlation pages

A *4D reconstruction* page that shows the resulting virtual reconstruction in an interactive way, with links to the related hypothesis pages. If an evolution in time is available, the model can be 4D. Techniques such as Flash of QuickTime VR can visualise such 4D models.

Finally, access to all this pages are provided through one or more *overview pages* that give access to the hypotheses for the object to visualise. An overview page can contain multiple objects (for example a list of buildings) and overview pages can be organised in a tree like structure for large and complex sites. As primary and secondary source pages, source correlation pages and 3D/4D visualisation pages are referenced from hypothesis pages, we don’t list them explicitly on the overview page.

8. Benefits

This methodology has several benefits for the different stakeholders involved in a 3D visualisation process.

First of all, as there is very little standardisation in how to conduct and document 3D visualisation research, this methodology helps to *structure and rationalise the interpretation process.* Currently, the interpretation process behind a 3D visualisation project is in most cases a black box with certain inputs and outputs but very little transparency concerning the process itself. Using some commonly accepted methodology will be beneficial for mastering the process and its quality.

Secondly, by recording the interpretation process through an online tool, other scholars or 3D visualisation specialists can understand the process and contribute their knowledge, through the known wiki mechanisms of discussion and consensus. This creates not only *scientific transparency* but stimulates also *multidisciplinary cooperation* as specialists in certain domains (for example stability analysis or building historians, specialised in a certain era) can be invited easily to contribute.

In other words, the proposed tool provides a *collaboration platform* to bring together all necessary specialists around the research and/or public presentation through 3D visualisation of historical manmade structures or landscapes.

By hosting this tool on some central server, managed by a central cultural heritage organisation in every country or region, all 3D visualisation processes can be *recorded and stored*, while the organisation itself can take care of all backup and long term storage, including all software updating and data migration in a user transparant way.

As most 3D visualisation projects are funded by public money, a supplementary requirement to record the corresponding interpretation process through such a centralised tool would yield not only a *long term storage* of knowledge that would otherwise disappear, a better safeguarding of the financial and intellectual effort that went into 3D visualisation projects, but also *general availability of 3D visualisation results* for the related community and for resuse in other projects.

Whenever new or updated information becomes available, the underlining database of the tool can be searched and all projects that use that specific information can be earmarked for update. Specialists can be invited to work on such an update or simply a list of projects that need update could invite specialists to donate time to integrate these new or updated results into the 3D visualisations. In the same way, results that would be reused will be earmarked for update, so no outdated 3D visualisations will be used or distributed.

9. Conclusion

The focus of 3D visualisation of historical structures is not 3D modelling or creating stunning images but conducting an indepth, systematic study of the sources, correlate and assess them, derive the most probable hypotheses, document this interpretation process in a well structured way and finally visualise them according the requirements of the context in which these visualisation results are used.

This paper provides a methodology that is on one hand flexible and capable of dealing with a wide range of subjects and goals, but on the other hand provides a standardised workflow which tries to turn 3D visualisation of historical structures into a repeatable, documented process that is transparent and publicly available.

In other words, this methodology for interpretation management establishes a sound framework for creating and publishing 3D visualisation results, improve their quality and preserve the investments and intellectual effort that has been spent to create them.

A specific EPOCH tool, based on wiki technology, has been realised to support this process and guarantee the safeguarding of the resulting data [EPOCH].

This paper is also available as an illustrated knowhow booklet for cultural heritage specialists [KHB08] and as a technical paper [PLE08], with many examples and a case study, on the EPOCH website [EPOCH].

Acknowledgements

I want to thanks first of all Marie-Claire Van der Donckt, curator of the Provincial Archaeological Museum Ename and Dirk Callebaut of the Flemish Heritage Institute, who directed the excavations at Ename, for providing me the opportunity to endeavour in the 3D visualisation effort of Ename from 1020 to 1780, in which I learned a lot about the issues and methodologies for 3D visualisation. These visualisations continue to play an important role in the Provincial Archaeological Museum in Ename.

I want to thank Anna Bentkowska-Kafel, Hugh Denard and Drew Baker of University College London who provided very valuable feedback on the concept and text of this paper. Together with Prof. Richard Beacham, Prof. Franco Niccolucci and many other 3D visualisation specialists, they are the driving forces behind the London Charter that plays a crucial role in setting standards and quality measures for 3D visualisation.

I want to thank Peter Sterckx to provide me his thesis at the KU Leuven on the 3D visualisation of the Horst castle in Belgium that shows the practical implementation and feasibility of interpretation management for large-scale projects.

I want to thank Joyce Wittur of the University of Heidelberg, who provided access in English to her ongoing PhD work and who took the effort and time to discuss the 3D visualisation process and the practical implementations within the cultural heritage community.

I want to thank Luc Van Gool to give me the confidence and support to realise the Interpretation Management tool under the EPOCH Common Infrastructure workpackage.

This paper has been realised with support by the European Commision under the Community's Sixth Framework Programme, contract no. 2002-507382.

References

[BOU00] Sophie Boukhari, Computers rebuild the Past, UNESCO Courier, online at http://www.unesco.org/courier/2000_03/uk/ signes/intro.htm, last verified on Feb 1, 2008

[EPOCH] EPOCH Network of Excellence, http://www.epoch.eu/, last verified on Feb 1, 2008

[FRI02] B. Frischer, F. Niccolucci, N. S. Ryan and J. A. Barcelo: *From CVR to CVRO: The Past, Present, and Future of Cultural Virtual Reality*. In: Franco Niccolucci: Virtual Archaeology: Proceedings of the VAST Euroconference, Arezzo 24-25 November 2000. Oxford, 2002. p. 7-18.

[JON07] Matt Jones, University of Southampton, Southampton in 1454: *A Three Dimensional Model of the Medieval Town*, 3DVisA Student Award Essay, http://3dvisa.cch.kcl.ac.uk/ student_award_3.html, last verified on Feb 1, 2008

[HAL03] Nick Halper, Mara Mellin, Christoph S. Herrmann, Volker Linneweber, and Thomas Strothotte: *Towards an Understanding of the Psychology of Non-Photorealistic Rendering*. In Jochen Schneider, Thomas Strothotte, and Winfried Marotzki, editors, Proc. Workshop Computational Visualistics, Media Informatics and Virtual Communities (April 4-5, 2003), pages 67-78, Wiesbaden, 2003. Deutscher Universitäts-Verlag.

[HER05] Sorin Hermon, Franco Niccolucci, Andrea D'Andrea, VAST-Lab, PIN, University of Florence, Italy: *Some Evaluations of the Potential Impact of Virtual Reality on the Archaeological Scientific Research*, VSMM2005 Proc., p. 105-114

[KHB08] EPOCH knowhow books, http://www.tii.se/knowhow/, last verified on Feb 1, 2008

[LARIS] Lorsch Abbey Reconstruction and Information System, http://www.iwr.uni-heidelberg.de/groups/ngg/Lorsch/ last verified on Feb 1, 2008

[NUME] Nuovo Museo Elettronico (NuME) - 4D visualisation of Bologna, http://www.storiaeinformatica.it/nume/ english/ntitolo_eng.html, last verified on Feb 1, 2008

[PLE08] Daniel Pletinckx, *An EPOCH Common Infrastructure Tool for Interpretation Management*, EPOCH 3.3 report, 2008

[QTVR] QuickTime VR, http://www.apple.com/quicktime/ technologies/qtvr/, last verified on Feb 1, 2008

[ROME] Rome Reborn 1.0, http://www.romereborn.virginia.edu/, last verified on Feb 1, 2008

[RYA96] Nick Ryan, *Computer-based Visualisation of the Past: Technical 'Realism' and Historical Credibility*, in P. Main, T. Higgins, J. Lang (eds.), Imaging the Past: Electronic Imaging and Computer Graphics in Museums and Archaeology, pg. 95-108, no. 114 in Occasional Papers, The British Museum, London, November 1996

[RYA01] Nick Ryan, *Documenting and Validating Virtual Archaeology*, Archeologia e Calcolatori, 12, pg. 245-273, November 2001.

[STE07] Peter Sterckx, *Digitale reconstructie en architectuurhistorische studie van het kasteel van Horst, te Holsbeek*, Thesis KU Leuven, 2007

[TLC] The London Charter, http://www.londoncharter.org/, last verified on Feb 1, 2008

[VAN06] Han Vandevyvere, Herman Neuckermans, Krista De Jonge: *Digital historical reconstruction: Case studies of an interdisciplinary task*, International Journal of Design Sciences & Technology, Vol. 13, N° 1, 2006, pp 51-65, http://www2.asro.kuleuven.ac.be/asro/english /HOME/Hvdv/Reconstruction/Digital historical reconstructionFormatted.pdf , last verified on Feb 1, 2008

[VAT03] I. Vatanen: *Deconstructing the (Re)Constructed: Issues in Conceptualising the Annotation of Archaeological Virtual Realities*. In: M. Doerr and A. Sarris: CAA 2002: The Digital Heritage of Archaeology: Computer Applications and Quantitative Methods in Archaeology. 2003. p. 69-74.

[WIT08] Joyce Wittur, University of Heidelberg, *Argumentation Networks*, Seminar "3D-Virtualisierung am Beispeil des Klosters Lorsch", 2007, http://www.iwr.uni-heidelberg.de/groups/ngg/ Lorsch/aims.php?L=E , last verified on Feb 1, 2008

3 EPOCH tools: features and availability

This section gives a short overview of the Epoch tools, one page per tool. Apart from a short description, it lists for each tool the following features:

- availability of the tool or its accessibility, also after the end of the Epoch project;
- conditions to use the tool;
- supported I/O formats;
- available documentation;
- contact information.

CityEngine

The CityEngine is a complete software suite, for the fast creation of 3D building models. Following the principles of procedural modeling, it generates compact models on the basis of the sequential application of a series of grammatical rules. Each architectural style comes with its own rule set. Starting from a crude, volumetric description, subsequent rules gradually expand the model obtained so far by replacing structures by more detailed ones. The resulting models are highly structured and compact, and the visual realism which can be achieved at limited time investment is unrivaled. The CityEngine comes with an extensive user interface, which allows for swift interactive changes.

CityEngine – ETH Zürich

Availability of the tool

The CityEngine software will be commercialised by Procedural, a spin-off company of ETH Zürich.

Conditions:

The pricing / business-model: software licenses, 7 kEuro per license for commercial use, 2 kEuro for academic partners
Commercial Epoch-members: discount of 20-30%

Supported Formats:

- GIS: KML, ESRIs Shape and DTMs (maybe also City-GML)
- 3D: Wavefront Obj, DXF, Maya, RenderMan, Mental Ray, Collada, FBX (Autodesk), OpenSG (maybe: 3DS (3D Studio), U3D (Acrobat),XSI (Softimage), X3D, FLT, PLY)

Documentation:

Information is available from the website of Procedural.

Contact:

Website: http://www.procedural.com
E-mail: info@procedural.com

EPOCH

MiniDome – K.U.Leuven

MiniDome

This apparatus is a hemispherical structure, with 264 white power LEDs, under computer control. There is one overhead camera, focused on the center of this dome structure. The dome itself consists of 4 shells, which can be easily assembled or disassembled. For the moment, the dome yields IBR results on the basis of estimates of local surface orientation. Cuneiform tablets or coins are good examples for the kind of objects we intend to digitise with the system, and we already collaborate with the assyriologists at the universities of Leuven and Cornell.




EPOCH

MiniDome – K.U.Leuven

Availability of the tool

The tool is a single integrated system, with all required hardware and software. The hardware can be transported easily (dome can be disassembled)

Accessibility after Epoch:
Copy of the mini-dome is under construction at K.U.Leuven, not financed by Epoch, for future use in Cultural Heritage

Conditions:
preferably rental service through Network of Expertise Centers
daily fee of 50 Euro, to guarantee swift turn-around times, transportation costs also for user

Supported formats:
input are objects...
output: proprietary format for the moment, OpenSG visualisation under study

Documentation:
Will be available with delivery of the MiniDome

Contact:
Website: http://www.esat.kuleuven.be/psi/visics
E-mail: Luc Van Gool: Luc.VanGool@esat.kuleuven.be

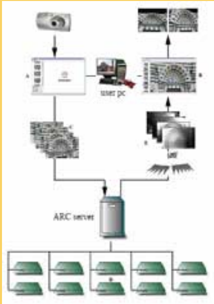
EPOCH

ARC3D (Automated Reconstruction Conduit) – K.U.Leuven

ARC3D

Users upload images of a scene or object to the ARC service over the Internet. A 3D reconstruction is produced fully automatically. Once the result is ready, the user is notified via e-mail that she can download the result, via ftp. The result consists of the 3D information as well as the relative vantage points from where the uploaded images were taken and the settings of the camera(s). The 3D information can be downloaded in any of a series of popular data formats.

SYSTEM: The upload tool (A) is responsible for uploading the actual input (C) to the ARC server . It supports the user in choosing the images to be used by the server, and possibly in further selecting the parts therein that are to be reconstructed in 3D. It is worthwhile spending some time on preparing the input, as this will result in a higherquality, cleaner result. In order to serve multiple users as quickly as possible, the server distributes the work over a network with hundreds of processors (D). Once the image sequence has been processed, the user can download the result, consisting of a depth map, a texture, and camera parameters per image (E). The viewer application (B) allows the user to look at the 3D model from different viewpoints, with or without texture and to save it in a series of popular data formats. ARC is compatible with Meshlab, an open source suite of tools to clean, simplify or complete the resulting 3D reconstructions.



EPOCH

ARC3D (Automated Reconstruction Conduit) – K.U.Leuven

Availability of the tool

The server running ARC3D at K.U.Leuven will be kept up. The service will be continued as long as personnel will be available to support it.

Conditions:

free for non-commercial use
license-based for commercial use, price on a per-case basis
Favorable conditions for Epoch commercial partners, to be discussed

Supported formats:

input: JPEG (also internally used), TIFF, BMP, PNG
output: native: VRML, OpenSG, OpenInventor, OBJ, PLY (MeshLab), X3D additional formats with MeshLab exporter (preferred way of use)

Documentation:
Documentation, tutorials, and examples are available from the ARC3D website.

Contact:
Website: http://www.arc3d.be
E-mail: Luc Van Gool: Luc.VanGool@esat.kuleuven.be

The Epoch Viewer – CGV, TU Graz

EPOCH

Epoch Viewer

The Epoch Viewer is a tool for interactive display and exploration of 3D scenes containing digital 3D artifacts. It is not only a viewer for static, pre-defined scenes, but it also allows to compose, edit, and save 3D scenes. It supports 3D object formats that are specifically developed for CH. It has an integrated scripting language that allows to attach scripted behaviour to each object, which can even be used to realize 3D user interfaces.

Technically the Epoch Viewer builds upon the open-source OpenSG scene graph engine together with the Generative Modeling Language (GML) as scripting language. The scene graph is basically a tree of nodes, where the root node represents the world coordinate system, and child nodes are typically transformations or geometric objects. Objects can be loaded from 3D files or from internet URLs (in various formats) or are created on the fly procedurally from a few high-level parameters (using GML).

The Epoch Viewer is not designed as a conventional stand-alone application, but it is rather a software component that consists of a 3D window. It can communicate with other applications that send commands or request information (using sockets). This makes it very flexible and easy to integrate it with existing applications. Of course, it is also easy to provide the Epoch Viewer with a conventional GUI that most users are familiar with.

The Epoch Viewer – CGV, TU Graz

EPOCH

Availability of the tool

The Epoch Viewer is maintained by the CGV group of TU Graz. It will be continuously developed and extended also after the end of Epoch. Besides its use in the CH sector it will serve as research platform to bridge the gap between semantic technologies (e.g., CIDOC-CRM) and 3D technology. Future work will range, e.g., from the semantic browsing of 3D digital libraries to the authoring of virtual test scenarios in cognitive psychology.

Conditions:
The Epoch Viewer is freely available for non-commercial use. For the regular use in museum exhibitions, we can offer very attractive conditions, depending on the requested level of support.

Supported formats:

- The set of 3D formats supported by OpenSG (VRML, .obj, .ply and others)
- The Collada Light format for scene graphs augmented with semantic information
- The Nexus file format for massive multi-resolution meshes from CNR Pisa
- The BTF file format for high-quality surfaces with complex reflectance characteristics

Documentation:
Documentation, tutorials, and examples are available from the Epoch Viewer website.

Contact:
Website: http://www.cgv.tugraz.at/EpochViewer
E-mail: Sven Havemann: s.havemann@cgv.tugraz.at
Institute of ComputerGraphics and KnowledgeVisualisation, TU Graz, Austria

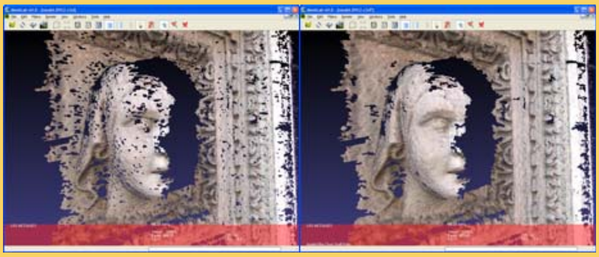
EPOCH

MeshLab – ISTI - CNR

MeshLab

MeshLab is an open source, portable, and extendible system for the inspecting, processing, editing and converting of unstructured 3D triangular meshes. It supports a wide variety of 3D formats and allows sophisticated automatic cleaning and repairing tasks.

Typical use: process and clean and convert 3D meshes obtained from 3D scanning technologies for Cultural Heritage purposes



EPOCH

MeshLab – ISTI - CNR

Availability of the tool

Thanks to its vast audience (thousands of active users) MeshLab has an active open source community that actively contributes to the base code. It will supported and expanded in the next years.

Conditions:

MeshLab is an open source tool that is distributed under the GPL2 licensing scheme.

Supported formats:

Input: PLY, STL, OFF, OBJ, 3DS, COLLADA, PTX
Output: PLY, STL, OFF, OBJ, 3DS, COLLADA, VRML, U3D, DXF

Documentation:

A wiki based documentation with some simple tutorials and reference info
A mailing list, and a tracker based public bug repository

Contact:

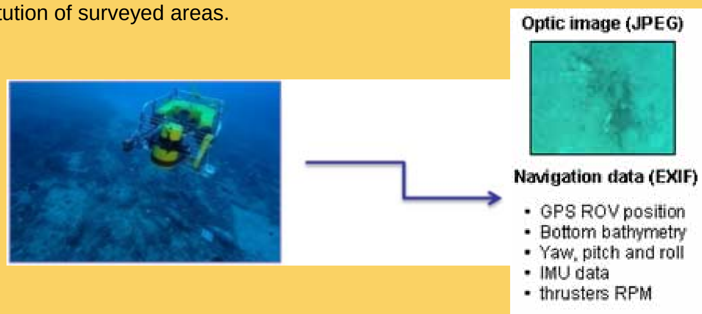
Website: meshlab.sourceforge.net
E-mail: p.cignoni@isti.cnr.it

EPOCH

EPL – ISME

EPL – Exif Postcard Library

Software Virtual Instruments for generating georeferenced, enriched visual data from underwater Photo and Video.
This software tool will generate EXIF postcards by associating to photos and video frames collected by underwater robots acoustic, geographic and navigation data, acquired by the NGC system of the robot.
Software library can be included into ROV and AUV NGC systems for JPEG/EXIFs restitution of surveyed areas.



EPOCH

EPL – ISME

Availability of the tool

The tool will be available through ISME website, Italy.

Conditions:
EPL will be freely available upon request. Details can be found on the EPL website.

Supported formats:
Libraries in LabVIEW, standard ANSI C calling convention dll.
INPUT: Hardware DATA coming from sensors.
OUTPUT: JPEG/EXIF

Documentation:
http://www.isme.unige.it/ /tools/epl

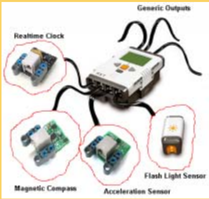
Contact:
Website: http://www.isme.unige.it/tools/epl
E-mail: labmacs@diiga.univpm.it;

EPOCH

LEGO Tracker – ISME

LEGO Tracker

Low cost device for underwater positioning and tracking of Divers. The device consists of a portable set of sensors, controlled by a microcomputer, for acquiring geographic and attitude data to complement photo and video surveys made by divers. Device construction is based on COTS and rapid prototyping methods.



EPOCH

LEGO Tracker – ISME

Availability of the tool

The tool will be available through ISME website, Italy.

Conditions:
Description, design and instruction for replicating the device will be publicly available. Details can be found on the LEGO tracker website.

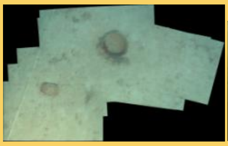
Supported formats:
Output of the device is a log standard text file downloadable by Bluetooth or USB connection with LEGO standard tools.

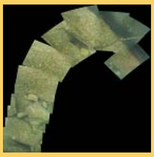
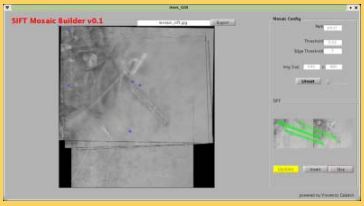
Documentation:
http://www.isme.unige.it/ /tools/LEGOtracker

Contact:
Website: http://www.isme.unige.it/ /tools/LEGOtracker
E-mail: labmacs@diiga.univpm.it;

SMuS - Snake Mosaicing using SIFT

Software suite for real-time photomosaicing from underwater Photos and VideoFrames. The software suite with user interface for rapid photomosaicing from underwater photos and video frames of a surveyed area.



SMuS – ISME

Availability of the tool

The tool will be available through ISME website, Italy.

Conditions:
SMuS will be freely available upon request. Details can be found on the SMuS website.

Supported formats:
Suite in MatLAB.
INPUT: JPEG/EXIF images
OUTPUT: JPEG/EXIF images.

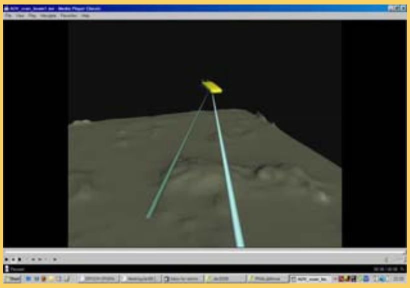
Documentation:
http://www.isme.unige.it/ /tools/SMuS

Contact:
Website: http://www.isme.unige.it/tools/SMuS
E-mail: labmacs@diiga.univpm.it;

EPOCH

MVT – Marine Visualization Toolkit

MVT provides marine cultural heritage professionals with the necessary tools to easily visualize and interact with bathymetric and other terrain datasets. It also provides pilots of Remotely Operated Vehicles (ROVs) with the means to plan and visualize dives either offline or in real-time scenarios.



EPOCH

MVT – University of Hull

Availability of the tool

The tool will be available through MARINEVIS website, UK.

Conditions:
MVT will be freely available upon request. Details can be found on the MVT website.

Supported formats:
All standard 3D formats accepted (HTML, 3DS, MATLAB etc). See documentation for full list. Real-time visualization for ROV positioning (if required) is achieved via TCP/IP.

Documentation:
http://www.marinevis.com/tools/mvt

Contact:
Website: http://www.marinevis.com/tools/mvt
E-mail: paul.chapman@hull.ac.uk

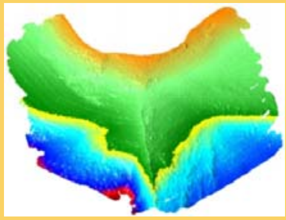
EPOCH

TerMap – IST

TerMap – Terrain Mapping Hardware and Software Suites

Mechanically scanning pencil beam-based system for bathymetric surveys in shallow waters, consisting of hardware and software suites for bathymetric data acquisition and processing. The tool allows for terrain map generation for archaeological applications using affordable systems.

Echosounder Image



Terrain Map

EPOCH

TerMap – IST

Availability of the tool

The tool will be available through DSOR/ISR/IST website, Portugal.

Conditions:
Tool description and general instructions for its use will be publicly available via the TerMap website. Upon request, consultancy on how to replicate and use the tool will be provided.

Supported formats:
Libraries in C++ and Matlab.
INPUT: Raw Data log files coming from acoustic sensors and attitude /position sensors
OUTPUT: Georeferenced points (ASCII XYZ , Matlab MAT, DXF), terrain surface model (VRML), Elevation maps (geoTIF)

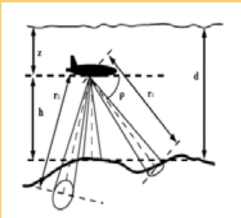
Documentation:
http://dsor.isr.ist.utl.pt/tools/TerMap

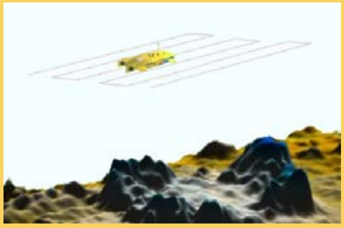
Contact:
Website: http://dsor.isr.ist.utl.pt/tools/TerMap
E-mail: antonio@isr.ist.utl.pt

EPOCH

TerNav – Acoustic Terrain-Based Navigation System

Algorithms and software suite for underwater platform positioning using ***acoustic terrain-based navigation****. The tool has the potential to be used for accurate ROV and AUV positioning over selected areas, the maps of which can be obtained using the TerMap tool. Specially suited for repeated operations over a site.*





EPOCH

TerNav – IST

Availability of the tool

The tool will be available through DSOR/ISR/IST website, Portugal.

Conditions:
Tool description and general instructions for its use will be publicly available via the TerNavwebsite. Upon request, consultancy on how to use the tool will be provided.

Supported formats:
Libraries in C++ and Matlab.

Documentation:
http://dsor.isr.ist.utl.pt/tools/TerNav

Contact:
Website: http://dsor.isr.ist.utl.pt/tools/TerMap
E-mail: antonio@isr.ist.utl.pt

EPOCH

AMA – PIN

AMA – Archive Mapper for Archaeology

AMA is a tool created for mapping existing archaeological datasets, including excavation databases, museum collections and free text data, to a CIDOC-CRM compliant format.

CIDOC-CRM is a standard ontology providing a rich and powerful data structure that simplifies the process of integration and the interoperability between existing and future archives. Information encoded using the CIDOC-CRM model is ready to be used in the forthcoming Semantic Web scenarios, extension and future of the present World Wide Web.

The AMA web tool allows users to import XML data models of their existing archaeological archives and to map them to the CIDOC-CRM ontology schema, creating direct relations with the CIDOC classes. The tool also allows the definition of simple and complex relations among the already mapped elements to enrich their semantic meaning.

The final result of the AMA mapping process is the creation of a mapping file to be used as a template for the data conversion process, directly in the original databases or together with conversion tools like MAD or D2RServer. Exported data will be ready for integration with other CIDOC-CRM compliant archives and for use in Semantic Web contexts.

EPOCH

AMA – PIN

Availability of the tool

The AMA online mapping tool will be accessible directly from the EPOCH website
http://www.epoch-net.org/AMA/tool/

Conditions:
AMA tool is an Open Source PHP on-line application. Source code will be available to download and use under the GPL license.

Supported formats:
Input: any XML compliant document, including semantic grammars (RDF, OWL) and schema formats (XSD, RDFS)
Output: mapping files in any XML compliant format, including XSLT

Documentation:
All the necessary documentation is available on the AMA web pages at http://www.epoch-net.org/AMA/

Contact:
Website: http://www.epoch-net.org/AMA/
E-mail: Andrea D'Andrea (dandrea@unior.it) - Achille Felicetti (achille.felicetti@pin.unifi.it)

EPOCH

MAD – PIN

MAD – Managing Archaeological Data

MAD is an application designed to store, manage and browse structured and unstructured archaeological datasets encoded in a semantic format. It can be used as an on-line tool or can be easily integrated with existing or newly developed applications.

It was created using Open Source technologies and is entirely based on XML and W3C standards, to act as the ideal framework for integrating and sharing data coming from diverse archives and encoded using ontologies such as the CIDOC-CRM.

MAD includes modules to encode entire datasets in a semantic format using an on-the-fly mapping mechanism driven by mapping templates (like the ones created using the AMA tool). Converted archives can be stored in the native XML database on which MAD is built, and immediately available for browsing and querying.

MAD provides a semantic browser with faceted browsing interfaces to access and filter information stored in the XML database in an innovative and meaningful way. MAD query features include simple, complex and semantic query. Basic geographic query features are also provided.

MAD modules are flexible and extensible enough to implement a semantic engine for any kind of context requiring semantic capabilities.

EPOCH

MAD – PIN

Availability of the tool

MAD will be released in 2 versions: an online application (http://www.epoch-net.org/MAD/tool/) and a set of modules downloadable from the EPOCH website (http://www.epoch-net.org/MAD/download/).

Conditions:

MAD tool is an Open Source Java application. Source code will be available to download and use under the GPL license.

Supported formats:

Input/output: any XML compliant document, including semantic grammars (RDF, OWL), schema formats (XSD, RDFS), stylesheets (XSLT), graphic formats (SVG), geographic formats (GML) and 3D formats (COLLADA/X3D)

Documentation:

All the necessary documentation is available on the MAD web pages at http://www.epoch-net.org/MAD/ and will be included in the downloadable package.

Contact:

Website: http://www.epoch-net.org/MAD/
E-mail: Achille Felicetti (achille.felicetti@pin.unifi.it)

Context Management Tools – Un. Kent & ARCES

MobiComp and Cimad Application Broker

- These tools provide components to support manage and coordinate context-aware applications in the Cultural Heritage domain.
- The main component is *MobiComp*, a framework to store, retrieve and aggregate *context elements*, i.e. heterogeneous static and dynamic information about the entities involved in an application (e.g. people, exhibits, devices, sensors). A *context-element* includes: a subject-predicate-object triple - relating an entity identifier to a named context value - a time-stamp, and additional specifiers (e.g. validity period and privacy level).
- *MobiComp* is a Java package.
- *MMPI* - MobiComp Multi Platform Interface - is an alternative - servlet based - Interface to MobiComp that removes the need to have a JVM installed on the client side. With *MMPI* MobiComp may be accessed from "any" platform (language, OS) provided with a network interface and a browser.
- *CAB (Cimad Application Broker)* is an *MMPI* based PHP web site. *CAB* registers on MobiComp Context-aware applications published on the web (*CIMAD applications*) and enables the users to choose their preferred CIMAD application.

EPOCH

Context Management Tools – Un. Kent & ARCES

Availability

MobiComp, MMPI and CAB will be available from the EPOCH website, til Dec 31, 2009, as software, in the following formats:

- *MobiComp:* Java source code
- *MMPI*: PHP and Java source code
- *CAB*: PHP source code

Conditions:

Agreed EPOCH tools Open Source Licence, free for EPOCH partners.
Users should be aware that these areconcept/early-stage tools.

Input formats:

N/A

Technical Documentation:

Software and documentation will be available from the MobiComp wiki at http://www.mobicomp.org/mediawiki

Contacts:

Nick Ryan, Computing Laboratory - University of Kent, Canterbury CT2 7NF, UK, n.s.ryan@kent.ac.uk – Phone: +44 1227 827699

Daniele Manzaroli, ARCES, University of Bologna - I
dmanzaroli@arces.unibo.it - Phone: +39 051 209 5426

Address: Via Toffano, 2 - 40125 Bologna - I

Path Finder

- This tool helps the visitors of cultural heritage sites to find their way towards a desired destination (a specific exhibit, an emergency exit, the book shop, etc…).
- *Path Finder* finds the "shortest" path to get to a desired location, starting from a know location. If used in conjunction with a location system, *Path Finder* may dynamically provide the direction to follow.
- *Path Finder* can be installed on a Web server (*Path Finder-WS*) or can be incorporated within a mobile application (*Path Finder-EMB*)
- *Path Finder* searches its path on a data structure (specifically a graph) representing the entire area covered.
- The graph may be generated by an off-line *Path Finder Remapping Utility,* starting from a site map coded in a specific bitmap format (see next slide)
- *Path Finder* may be delivered with its *Path Finder Remapping Utility,* and with a desk-top demo application.

PATH FINDER – ARCES

Availability

- *Path Finder* will be available from the EPOCH website, til Dec 31, 2009, as software, in the following formats:
  - *Path Finder* run time module*:* Windows Mobile DLL or Web Service (binary or source)
  - *Mapping Utility*: Windows based stand-alone binary application ((this utility is based on *OpenCV* library)
  - *Demo application:* Windows based stand-alone binary application

Conditions:

Agreed EPOCH tools Open Source Licence, free for EPOCH partners.
Users should be aware that these are concept/early-stage tools

Input formats:

A .BMP site map, where pedestrian areas are red painted

Technical Documentation:

Software and documentation will be available from the MobiComp wiki at http://www.mobicomp.org/mediawiki

Contact:

Daniele Manzaroli, PhD student at ARCES, University of Bologna
dmanzaroli@arces.unibo.it - Phone: +39 051 209 5426
Address: Via Toffano, 2 - 40125 Bologna - I

Appendix

Work on tools since last review

This appendix gives a short overview of the work on tool development that has been carried out since the last review.

AMA – PIN

The AMA tool has become a fully featured web application implemented in PHP and accessible from anywhere. New features have also been added to the preceding stand-alone version, including:

- The possibility to upload both starting schema and target ontology for any kind of mapping (not only towards CIDOC-CRM)
- Ontology to ontology mapping capabilities
- Advanced mapping features for complex mapping definitions (i.e. creation of new entities and properties to represent implicit elements and relations)
- Creation of shortcuts to simplify the mapping process
- Graphic visualization of simple and complex relations obtained during the mapping process.

ARC3D (Automated Reconstruction Conduit) – K.U.Leuven

- Further experimentation in connection with CNR's PhotoCloud

CityEngine – ETH Zürich

- Improved interface, for enhanced on-line response
- Reworked structure of the grammatical rules

Context Management Tools – Un. Kent & ARCES

- The main effort after last review was addressed to understand interoperability issues, advancing towards an ideal picture of seamless Common Infrastructure based services offered to every type of User Equipment, without the need of device customization. This activity was carried out in cooperation with other WP3 partners (mainly KUL, Polimi, PIN and Ducati).

- Interoperability was considered both from the Content and from the Execution Environment point of view.
- Interoperability at the Content level was pursued using Fedora CMS, as established by the CI coordinators
- Interoperability at the execution level was based on the principle that every application should be incorporated within a web page.
- Development activities where then focused on MMPI and CAB

The proposed tool-suite and the related architecture were presented at ICHIM07 in Toronto (Nov. 2007)

- G. Raffa, P. H. Mohr, N. S. Ryan, D. Manzaroli, M. Pettinari, L. Roffia, S. Bartolini, L. Sklenar, F. Garzotto, P. Paolini, T. Salmon Cinotti (2007), CIMAD - A Framework for the Development of Context-aware and Multi-channel Cultural Heritage Services, - ICHIM07 (International Cultural Heritage Informatics Meeting), Oct 24-26, 2007, Toronto

EPL – ISME

- The tool and a specific protocol to associate (in JPEG/EXIF format) navigation data to images taken by digital cameras or video cameras mounted on underwater vehicles have been tested in simulation environment (using part of images coming by other missions project) and in experimental activity (Tremiti Islands mission and Portugal partially supported UPGRADE Mission).

Epoch Viewer – CGV, TU Graz

- Socket communication and command interface that allows to control the viewer in all aspects from another application: Camera control, loading scenes and objects, manipulating the scene graph, issuing events, and requesting information about the scene and its objects
- Stereoscopic viewing in side-by-side mode, in two-window mode, and in quad-buffer mode to support the most widely used stereo setups
- Head tracking for single users: The viewing position is continuously updated to match the real-world position of the user. The user has the impression that the objects are flowing in mid-air, the screen is like a window to a virtual 3D world ("1-sided CAVE")
- Generalized input device interface, now also using sockets. Protocols were defined for the Wii, Joystick, and Spacemouse, as well as for 6-DOF devices whose real world pose (position and orientation) is determined by our camera-based tracking software. This is a low-cost, high-quality solution for 6-DOF tracking.

LEGO Tracker – ISMR

- The tool and a specific protocol have been tested with divers in experimental activity (Portugal partially supported UPGRADE Mission).

MAD – PIN

The MAD application has been enriched with a new set of query interfaces and new capabilities for storing and browsing semantic documents. New geographic features have been implemented, including:

- A module for indexing spatial data (GML) to be used in semantic contexts
- A set of XQuery spatial functions to query GML features
- A set of Java features to use AMA mapping files for on-the-fly publishing of databases in a semantic format.

MeshLab – ISTI - CNR

- The latest version 1.1.0 has been released in January 2008. Downloaded 2000 times in less than a week, it includes a new sophisticated interface for range map alignment, range map integration algorithms, and tools for measurement and creation of slices and sections.
- Another notable improvement is the possibility of converting 3D object in a format that is directly embeddable into a PDF and that can be interactively viewed (rotated and zoomed) directly inside acrobat reader.

MiniDome – K.U.Leuven

- Further work on filter sets, to enhance the legibility of marks, e.g. cuneiform writing on clay tablets
- Improvements on the interface

MVT – University of Hull

- Fast bathymetric terrain rendering using contemporary GPU technology
- Seabed measurement and plane fitting
- Accurate Geo-referencing
- Remotely Operated Vehicle (ROV) and Autonomous Underwater Vehicle (AUVV) dive planning with
  - Avoidance areas
  - Areas of interest with automatic route planning
- Dynamic real-time ROV positioning
- Underwater current modelling
- Stereo visualization for virtual reality

PATH FINDER – ARCES

- Path Finder was almost entirely developed after last review, as an add-on to EPOCH CI and it can be exposed to its users by CAB (Cimad Applications Broker).
- Path Finder offers a service that mobile applications for museums and cultural heritage sites may require.

SMuS – ISME

- The tool and a specific protocol to merge images taken by digital cameras or video cameras mounted on underwater vehicles have been tested in simulation environment (using part of images coming by other missions' project) and in experimental activity (Tremiti Islands mission and Portugal partially supported UPGRADE Mission).

TerMap – IST

IST completed the work on and tested a mechanically scanning pencil beam-based system for bathymetric surveys in shallow waters.

- Hardware: single beam echosounder, GPS with RTK, attitude and reference unit. Can be installed on-board a robotic marine vehicle or a manned surface vessel.
- Software for data acquisition and processing (including outlier detection and geo-referencing) and map construction.
- Tests carried out in Sines, Portugal confirmed the utility of the tool developed.

TerNav – IST

- IST finalized the development of advanced algorithms for terrain based navigation using acoustic sensor suites.
- Extensive Monte Carlo simulations were done to assess the efficacy of the tool developed.