VSMM 2008

Digital Heritage

Proceedings of the 14th International Conference on Virtual Systems and Multimedia

Full Papers

20–25 October 2008
Limassol, Cyprus

M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis (Editors)

VSMM 2008 – Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia

This volume contains the Research Papers presented at VSMM 2008, the 14th International Conference on Virtual Systems and Multimedia which took place on the 20 to 25 October 2008 in Limassol, Cyprus. The conference title was "Digital Heritage: Our Hi-tech-STORY for the Future, Technologies to Document, Preserve, Communicate and Prevent the Destruction of our Fragile Cultural Heritage".

The conference was jointly organized by CIPA, the International ICOMOS Committee on Heritage Documentation and the Cyprus Institute. It also hosted the 38th CIPA Workshop dedicated on e-Documentation and Standardization in Cultural Heritage and the second EuroMed Conference on IT in Cultural Heritage. Through the Cyprus Institute, VSMM 2008 received the support of the Government of Cyprus and the European Commission and it was held under the Patronage of H. E. the President of the Republic of Cyprus.

The Cyprus Institute

International Council on Monuments and Sites
Conseil International des Monuments et des Sites

CIPA Heritage Documentation

ΤΕΧΝΟΛΟΓΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ 2004

UNESCO United Nations Educational, Scientific and Cultural Organization
With the support of UNESCO's Culture Sector

isprs

VSMM 2008

Digital Heritage

Proceedings of the 14th International Conference on Virtual Systems and Multimedia

Full Papers

20–25 October 2008
Limassol, Cyprus

M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis (Editors)

Marinos Ioannides
Editor-in-Chief

Elizabeth Jerem
Managing Editor

Fruzsina Cseh, Elizabeth Jerem
Copy Editors

ARCHAEOLINGUA
Cover Design

Cover image: 3D reconstruction of faces appearing in Cultural Heritage Artifacts. The top row shows actual images showing the faces of Saint Mamas and Saint Tychikos from the wallpaintings in the Church of Panayia Phorbiotissa - Asinou in Cyprus (UNESCO World Heritage Site: http://whc.unesco.org/en/list/351). The remaining rows show the corresponding reconstructed 3D models as seen from different viewpoints. More information about the 3D reconstruction method can be found in the research paper “Reconstructing 3D Faces in Cultural Heritage Applications’’ by A. Lanitis and G. Stylianou.

ISBN 978-963-8046-99-4

Published by ARCHAEOLINGUA
Printed in Hungary by PRIMERATE

Budapest 2008

ARCHAEOLINGUA

VSMM 2008

Digital Heritage
Proceedings of the 14th International Conference on Virtual Systems and Multimedia

20–25 October 2008

LIMASSOL, CYPRUS

Under the Patronage of H.E. President of the Republic of Cyprus

Cyprus Government

In cooperation with

Cyprus Government

Department of Antiquities in Cyprus

In cooperation with European Union Projects

CHIRON

COINS

Institutional Sponsors

Ministry of Education & Culture

University of Cyprus

Supporters

The official carrier of the joint event

Acknowledgements and Disclaimer

The VSMM 2008 joint conference has been partly supported by the VSMM Society, the Cyprus Government, the Cyprus Institute, The Cyprus University of Technology and the University of Cyprus.

The 38th CIPA International Workshop has been supported by CIPA, ISPRS and ICOMOS.

The 2nd EuroMed Conference has been supported by UNESCO-Cyprus Committee and the Cyprus Government.

Foreword

These conference proceedings contain a selection of papers that focus on multi-disciplinary research involving both Cultural Heritage (CH) Informatics and also the use of technology for initial data-capture and digitization, information data-processing, reconstruction, modelling, visualization, documentation and archiving, as well as visualisation of results and dissemination to the scientific and cultural-heritage communities and to the public. The contributions in these proceedings will definitely assist all experts involved in Cultural Digital Heritage in restoring, renovating, protecting, documenting, archiving, and monitoring history and prehistory, to secure this information for years to come. It is clear that a worldwide collaboration in this area will help make the past accessible to the present and the future.

Cultural Heritage is being transformed by the nature of digital representation of culture in which production, documentation, and distribution of an artefact are one and the same. Understanding and defining digital cultural heritage has implications for documentation practices and the experience of cultural institutions.

Digital devices provide unique access to archives and cultural exhibits, enhancing the capacity of museums and collections to encourage community building and civic engagement. Collection databases once used solely by museum professionals are now being made available locally and globally through the Web. Increasingly, access to cultural heritage is digital and experienced through electronic images and facsimiles. Digital tools and information and communication technologies are merging as the basis for preserving cultural heritage. Digital 3D modeling provides precise and complete documentation of cultural heritage objects and sites and should be used in conjunction with traditional techniques. Of great interest to the scientific community in the last few years, especially in the areas of architecture and preservation, are 3-D modeling, visualization and animation of cultural-heritage monuments and sites. The cooperation between photogrammetry and computer graphics has led to the development of new tools and techniques that are particularly useful for the documentation and archiving of cultural heritage in a digital format. These new tools and techniques include not only photogrammetry, but also 3-D reconstruction, visualization, animation and virtual reality. Technical achievements in modeling, rendering, and animation have made possible the creation of virtual environments, providing a convincing visual experience of cultural heritage structures and sites.

The island of Cyprus is a particularly appropriate venue for a conference on Virtual Systems and Multimedia dedicated to CH because of the long and rich pre-history and history of the island. The historical and archaeological context of Cyprus is the physical and ideal setting of this conference. The past story of Cyprus is the history of the interaction of the cultures and peoples of the lands surrounding the Mediterranean Sea, which was the central means of communication, transport, trade and cultural exchange between diverse peoples. Its history is important to understanding the origin and development of the Mesopotamian, Egyptian, Persian, Phoenician, Jewish, Greek, Roman, Arab and Ottoman cultures and, hence, is important to understanding the development of Western civilization as we understand and experience it today.

The roots of cultures and civilizations are embodied in their architectural structures and archaeological sites, and this cultural heritage should be preserved for future generations. The importance of preservation can be seen in the efforts of international organizations to document important structures and sites. UNESCO and ICOMOS have called for all national and international organizations that are responsible for manmade monuments to document cultural-heritage objects and sites with methods that include traditional and newer, innovative technologies. The integration of these technologies offers great promise and the use of digital technology in particular has rapidly changed documentation techniques.

The importance of Digital Cultural Heritage is evident by the participation and cooperation of a large number of people and organizations including the following:

- The 14th International Conference on Virtual Systems and Multimedia VSMM 2008, dedicated on Digital Heritage (http://www.vsmm.net/)
- The 38th CIPA International Workshop dedicated on e-Documentation and Standardization in Cultural Heritage (http://cipa.icomos.org)
- The 2nd Euro-Med Conference on IT in Cultural Heritage.

We extend our thanks to all those, whose labour, financial support, and encouragement made this joint event possible. The International Program Committee, whose members represent a cross-section of Archaeology, Computer Graphics and Design, Architecture, Surveying, History and Engineering worked tenaciously and finished their work on time.

Especially Mr. Nikolas Valerkos, who designed and managed the webpage and Dr Andreas Lanitis who supervised the web-based submission system and guided the effort that published these proceedings. We would like also to express our gratitude to our co-organizers The Cyprus Institute, the Department of Antiquities in Cyprus and the Technological University of Cyprus. Finally, our institutional sponsors, the Ministry of Education and Culture, the University of Cyprus; and our official carrier, Cyprus Airways who provided money and 'gifts of kind' that made the conference possible.

Our Keynote Speakers, Javier Hernandez-Ros, European Commission, John Van Oudenaren, World Digital Library, *Library of Congress***,** Massimo Negri**,** Europeana and European Museum Forum; Prof. Donna J. Cox, *University of Illinois at Urbana-Champaign;* Vassilios Tsingas, Elliniki Photogrammetriki Ltd**;** Kareem M. Darwish and Ahmed El-Shimi, *Cairo Microsoft Innovation Center,* are not only experts in their fields but also visionaries for the future of IT in CH. They promote the e-documentation of the past in such a way for its preservation for the generations to come.

Most of all we would like to thank the Cyprus Government, the European Commission, UNESCO WHC, ISPRS, ICOMOS, VSMM-Society and CIPA, that entrusted us with the task of organizing and undertaking this unique event and wish all participants an interesting and fruitful experience.

Marinos, Alonzo, Andreas, Loukas
Limassol, Cyprus 2008

FULL PAPER INDEX BY SESSION

Data Acquisition and Remote Sensing in Cultural Heritage

INTEGRATION OF A TLS IN A LOW COST MOBILE MAPPING SYSTEM

M. Alshawa*, P. Grussenmeyer, E. Smigiel

Photogrammetry and Geomatics Group, MAP-PAGE UMR 694
INSA de Strasbourg, Graduate School of Science and Technology
24 Boulevard de la Victoire, 67084 STRASBOURG, France
(majd.alshawa, pierre.grussenmeyer, eddie.smigiel)@insa-strasbourg.fr

KEY WORDS: TLS, GPS/INS, navigation, integration, experiment, modelling, accuracy, mobile mapping

ABSTRACT:

Mobile mapping techniques have become a principal data source of 3D GIS, urban digital models and their applications using virtual and augmented reality. The high productivity of such systems involves usually high constructional cost. Thus, the constructors invest vastly in choosing high grade sensors and processors. Our choice, presented in this paper, is to simplify the design of mobile mapping in such a way that it becomes available for a large range of applications. A mobile mapping system of minimum requirement based on Terrestrial Laser Scanner (TLS) is presented. It can be mounted and operated by one person. The system is presented from an analytical point of view. Integration, synchronisation, and interpolation issues are detailed. The use of 3D TLS without any drastic modification is showed as a prototype for integrating any possible imaging sensor. The particularity of each component of the system is taken into account while preparing data to be integrated. Since the system is supposed to operate on whatever mobile platform, we present a fast method of precise system calibration.
Inertial navigation system augmentation techniques as magnetometers and zero velocity update do not involve high costs, so they are used to bridge eventual GPS outage. The effect of each mode of augmentation on the final point cloud accuracy is illustrated and compared.
The difference in geometric aspect between point clouds acquired in a stationary mode and those acquired by a mobile system yield a change in modelling procedures. The point cloud obtained after processing the data from the current configuration has an accuracy of ± 20cm. Thus, some related modelling hints are discussed to take into account the resolution and the accuracy.

1. INTRODUCTION

The progress in the world of satellite and inertial navigation allows further possibilities of kinematical surveying. One important application is the direct georeferencing which is the principal idea of mobile mapping. The term of mobile mapping could be extended to include remote sensing applications but it is often used to describe airborne and terrestrial mapping. The existence of airborne photogrammetry and LIDAR has preceded terrestrial application despite their common cost and principle of operating. This could be explained by the need to cover large areas by 2D maps. New requirements of full 3D spatial data have been created by 3D GIS, cultural heritage documentation, virtual and augmented reality. Hence, terrestrial mobile mapping revolution has accelerated, especially in the last decade.
Consecutive generations of mobile mapping platforms use more and more sensors. Georeferencing Sensors are basically an integration of Global Position System (GPS) and Inertial Measurement Unit (IMU). The GPS has been replaced by Global Navigation Satellite System GNSS to improve the constellation and to insure best visibility of satellites. Odometers and digital barometers are used also to bridge GPS outages and lack of precision in altitude respectively. Many algorithms are used to integrate all these sensors and to compensate the weak points of either.
The evolution of mapping has accelerated also with the increase of the spectral capacities of imaging sensor. The progress in microprocessors and wireless technologies reflected directly on mobile mapping systems especially in the synchronisation and real time processing aspects. So, mobile mapping platforms have become very sophisticated. They are either operated mostly by a team of various specialities or by a system integrator company. One can state that the mobile mapping is one of the most interdisciplinary work. All existing systems have not stabilized on a definitive configuration, but have a current composition of sensors as we will see in the next paragraph.
The aim of this paper is to present a concept of terrestrial mobile mapping system developed in the MAP-PAGE laboratory. Such experience could interest almost all laboratories using a terrestrial laser scanner (TLS). One of the basic ideas of this article is the possible duality fix/mobile of the TLS under certain constraints.
Firstly, the hardware and connexion rules are presented; then, necessary procedures to obtain the final point cloud are showed. Finally the features of the acquired point cloud by TLS in mobile mode are discussed.

2. RELATED WORK

First generation of mobile platforms has used cameras as imaging sensors (Goad 1991), (Tao et al. 2001). GPS/INS integration was used to provide the elements of external orientation of the photos obtained during the mobile sessions. Knowledge in close range photogrammety and navigation have been combined in order to generate 3D models for hundred of kilometres of entities adjacent to roads.
Another evolution has been introduced with the appearance of laser telemeter since 1999, as for example in (Gräfe, 2007). High accuracy and coverage was the main advantage to include these instruments on mobile platform. Used laser telemeters have to reach high speed of acquisition to enable mobile

platforms to run in normal traffic speed. Laser scanning and photogrammetry are used together because of there complementary features; photos provide the semantic appearance of 3D datum built from laser scanning. The integration between the two capturing devices can go further: (Gajdamowicz et al. 2007) has used the bundle adjustment to obtain the position and the attitude of the camera and consequently to recalculate the GPS/INS trajectography. This procedure is used finally to refine the precision of final point cloud acquired simultaneously.

Low cost mobile mapping systems are often based on digital cameras. The reason behind this use is not only their low price in comparison with a laser scanner, but the non gyroscopic attitude information obtained in a photogrammetric method. In such way, the use of inertial sensors of moderate performance will be feasible. Moreover, (Da Silva et al. 2003) depend only on photogrammetric operations with image pairs to orient the CCD cameras. The system contains a GPS antenna but no inertial unit. One of additional low coast aspects of this system is the synchronisation; GPS position recording time is characterised by a sound signal using the on-board laptop. The signal is sent to the video camera making a distinct noise which represent the GPS measurement time. (Madeira et al. 2007) established a mobile mapping system with two CCD progressive colour video cameras and a low cost navigation unit composed of a single frequency GPS, a MEMS gyroscope, and car odometer. The reached accuracy is about 1-2 m which is sufficient for detecting traffic signs for example.

Such experiences are rarely found in the systems using laser scanners. The absence of overlapping data in mobile laser scanning prevents the indirect georeferencing possibility. An external source is hence inevitable to accomplish the direct georeferencing.

The system discussed in this paper is not aimed to be compared with the high equipped multisensor ones, but with more simple systems which depend on a limited number of sensors at a low operatory cost (like the two systems described above).

Whatever the configuration of laser based terrestrial mobile mapping, it uses 2D telemeter laser: (Abuhadrous et al. 2004), (Hunter et al, 2006) for instance. In some exceptional case, such as (Kukko et al.2007), 3D TLS is used but after blocking its horizontal movement. Indeed no 3D function is used; the scanner is practically brought to 2D one.

3. SYSTEM COMPONENTS

The current configuration of the described system consists of three sensors (GPS, IMU and TLS). They are mounted on a simple carriage drawn by the speed of a walking person. The main kinematic and accuracy characteristics as follows:

- GPS Leica® GPS1200: working in differential mode. The used message is NMEA (GGA sentence for the position and VTG for the speed over ground). Sampling rate used was uniformed at 4 Hz. In the ideal observation condition (GDOP less than 4), GPS precision reaches 2cm in plan and 5cm in vertical. Nevertheless no quality control is used for real time connexion requirement. Some measures could have up to 2m accuracy;
- AHRS440 Crossbow®: Attitude and Heading reference System aided by GPS. It uses MEMS (Micro-Electro-Mechanical Systems) tri-axial magnetometer, sensor rates and accelerometer to provide 1.5° accuracy angle measurement. When the external GPS signal is available, the attitude accuracy improves up till 0.5° thanks to Kalman filter embedded on the AHRS micro processor.
- TLS: laser scanner 3D GX DR 200+ from Trimble®: the scan is performed vertically within 60° field of view. The maximum range of the scanner is 200m. It can work in non levelled (non-horizontal) mode. Maximum speed is reached when using a single laser shot for each measure. The last calibration yields a distance accuracy of 14 mm at 100 m (for one shot) and an angle accuracy of 12”-14” while measuring a single point.

Figure 1: The Trimble® GX scanner integrated on a mobile platform

The power of all system parts is supplied by batteries. Figure 2 shows how different components connect to each other. In view of the low displacement speed, the use of high sampling rate is not justified. Moreover the resolution of AHRS attitude is equal to 0.1°.

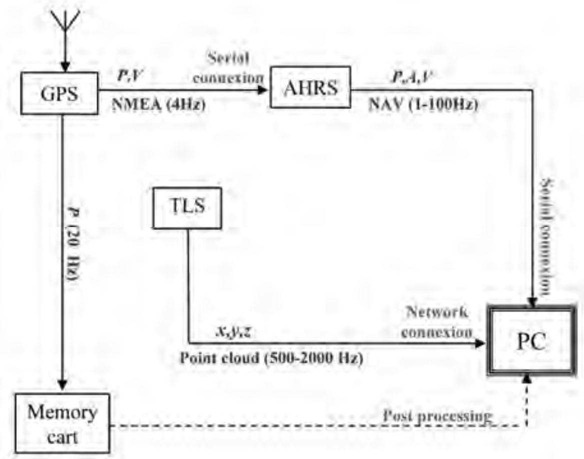


Figure 2: Connexion schema of different parts of the system; *P*: position, *V*: velocity, *A*: attitude, NAV: navigation message obtained by AHRS 440 (time, acceleration, angular rates, attitude, GPS coordinates, built in test indicator)

4. GPS/INS/TLS DATA INTEGRATION DESIGN

Data integration in our prototype is done in post-processing. After returning from mobile mission, we can observe three sources of spatial data:

- GPS coordinates recovered on GPS interior memory card (with a control of accuracy).
- Navigation messages logged into a text formatted file (no accuracy control is available)
- Point cloud as coordinates measured in the local frame of the scanner (the origin is always equal to zero and the north is the identical to the direction of scanner horizontal disc zero).

These data have to be involved into the following equation in order to obtain the final georeferenced point cloud (Alshawa et al. 2007):

$$r_p^g(t) = r_{IMU}^g(t) + R_b^g(t).[R_s^b . r_p^s(t) + a^b] \quad (1)$$

where:

r_p^g : Position of a point *P* in the geodesic frame;

r_p^s : Position of a point *P* in scanner local frame (given by TLS calculation module);

r_{IMU}^g : Position of IMU in geodesic frame (given principally by GPS at given frequency);

R_b^g : Rotation matrix from IMU (body) to geodesic frame (measured by AHRS at given frequency too);

R_s^b : Rotation matrix from scanner to IMU frame (to be calculated by prior calibration);

a^b : distance scanner-IMU (to be calculated by calibration too);

Equation (1) should be applied for each measure of TLS at the instant *t,* which imposes possessing values for each time dependent variable in the equation at the same instant *t*. Thus different data have to be interpolated and synchronised before processing them into the mathematical model. Time independent variables are calculated by calibration prior to processing. The next three paragraphs are devoted to operations previous to data integration by the model given by equation (1).

4.1 Data synchronisation

The synchronisation between GPS and the inertial measures is made by the AHRS microprocessor itself. Latency time passed between receiving GPS satellites signals and parsing NMEA code by the AHRS is assumed to be constant (20msec) for the AHRS calculating module. The difference between this constant delay and the real time needed to calculate GPS position and velocity and to dispatch them is the first source of synchronisation error. This error is quite negligible comparing to the second one explained later in this paragraph.

The normal use of a TLS does not involve any timing requirement, so scanned points have not any related date. The solution supposed here in is to intercept data flow by means of a network analyser and establish a relationship between the time of data acquisition and data entry.

For this purpose we use Wireshark which is a free network protocol analyzer. Data analysing is done simultaneously with data acquisition by PointScape®; the software which operates the scanner. Wireshark interface while parsing TLS data packets is illustrated in figure 3.

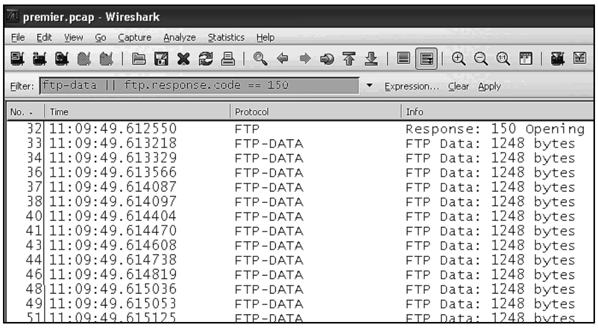


Figure 3: Communication between scanner and piloting computer as monitored by Wireshark

A certain systematic behaviour has been stated by noting a large number of samples of laser scanning. Indeed, the scanner sends a FTP (file transfer protocol) demanding the creation of a temporary file of known size. It starts then filling the file with binary data transferred by a FTP-DATA protocol. When the initial size of the file is reached, the scanner sends a FTP command to delete it and to create another one. This operation is repeated for each couple of scan line and takes about 0.01 sec. Thus, we can obtain only the final instant for each pair of scan lines.

A linear interpolation between two successive end times over the points measured in this interval seems to be the best estimation (certainly, time of transfer has to be taken in consideration). An instant is attributed for each point by this method. Nevertheless, there remain some reasons for estimation incertitude:

- The movement of the rotating mirror of the scanner is supposed regular.
- The horizontal movement between two vertical lines is supposed equivalent to the one between two points.
- The time of scanning the first couple of lines is estimated as the mean of other couple scan times.

In view of the method used for the synchronisation, the accuracy of the solution could be estimated as the time of one column of scan (about 0.075 sec while using the velocity maximal of the scanner). If the displacement velocity is equal to 1 m/sec, synchronisation error can cause a position error of 7.5 cm. Finally, we can state that synchronisation error has been minimised so far. However it could be reduced provided that one knows better the embedded electronics and control laws of the scanner (Trimble® GX). The systematic portion is not negligible and could be removed with several calibration tests.

4.2 Position and attitude interpolation

As explained above, the AHRS permits to couple GPS and data obtained from accelerometers, gyroscopes and magnetometers in order to improve the accuracy of calculated angles and position. A simple knowledge in navigation allows to know that the used algorithm is the Kalman filter with loosely coupled computing method. INS mechanisation yields an estimation which is affected by an error budget (mainly the gyro drift). Kalman filter prediction loop runs at the same sampling rate of the mechanisation using mathematical model to correct the calculated values. Once a valid GPS position is received, the correction loop runs in order to feed back the calculated correction to the output values from previous loops. Prediction- correction loops have not to run at the same frequency, nonetheless, position and attitude outputs have the same rate while no GPS outage takes place. While GPS and

INS data are coupled and synchronised, no interpolation is needed at this stage. On the other hand, the interpolation is an essential task when coupling GPS/INS and TLS data (as illustrated in figure 4.)

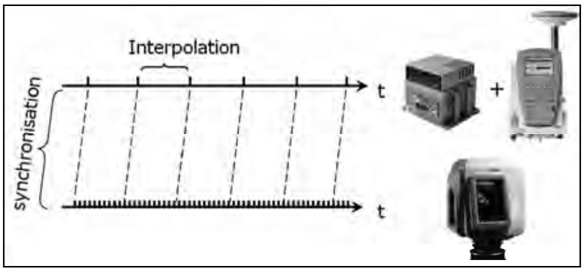


Figure 4: GPS/INS + TLS measure instants represented on a time scale

While the AHRS is based on MEMS sensors, it suffers from a high drift with time if no GPS signal is provided and can not consequently compute an accurate position. Long GPS signal absence causes thus a change in AHRS navigation message; no position value exists and attitude angles are less accurate. In this case, alternative calculation algorithms will be useful for stabilising the gyro drift by magnetometers measures.

One can state that the aim of the interpolation is to densify navigation measures to meet those of the scanner and to bridge GPS outages. In order to resolve the former, a simple linear interpolation is sufficient and curve fitting method is used for the latter.

(Nassar et al., 2007) advise using Kalman filter smoothing method to overcome GPS blockage. This method imposes a prior knowledge of error model used by the AHRS in real time integration. The use of the same model but in backward direction in post treatment ensure a higher accuracy while GPS blockage bridging.

While no information about Kalman filter used by AHRS is available, Polynomial curve fitting method is preferable at this stage of research. The method expresses each coordinate as n^{th} order polynomial function of time:

$$x = \sum_{j=0}^{k} a_j \, t^j \quad (2)$$

Where, a_j are the coefficients which are sought for.

For n measures, Vandermonde matrix V and vector b are formed as following:

$$V = \begin{pmatrix} n & \sum_{i=1}^{n} t_i & \cdots & \sum_{i=1}^{n} t_i^k \\ \sum_{i=1}^{n} t_i & \sum_{i=1}^{n} t_i^2 & \cdots & \sum_{i=1}^{n} t_i^{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} t_i^k & \sum_{i=1}^{n} t_i^{k+1} & \cdots & \sum_{i=1}^{n} t_i^{2k} \end{pmatrix} \quad b = \begin{pmatrix} \sum_{i=1}^{n} t_i \\ \sum_{i=1}^{n} t_i x_i \\ \vdots \\ \sum_{i=1}^{n} t^k{}_i x_i \end{pmatrix} \quad (3)$$

The least square solution of the formula $V.\hat{a} \cong b$ is given by:

$$\hat{a} = (V^T W V)^{-1} (V^T W b) \quad (4)$$

Where W is the weight matrix whose elements could be concluded coordinates quality control v (recovered from GPS memory support).

$$W = diag\left(1/v_1 \quad 1/v_2 \quad \cdots \quad 1/v_n\right) \quad (5)$$

After obtaining the coefficient values, one can substitute instants of each laser scanning point in equation (2) to calculate the related coordinate.

While all GPS coordinates are taken into account, whatever their precision, it is preferable to tolerate the conditions which constraint the acquisition (GDOP, Signal-to-noise ratio).

Figure 5 shows the result of horizontal interpolated coordinates using a polynomial of degree 13.

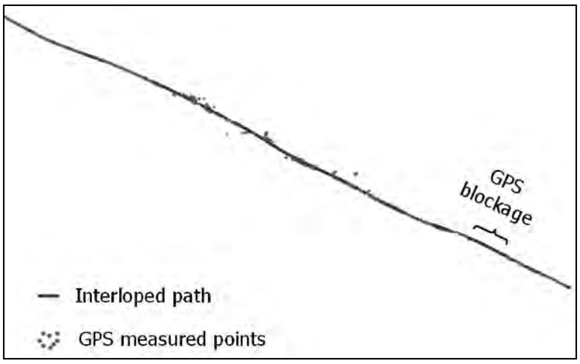


Figure 5: Bridging GPS blockages by interpolation

Though preferable scanning time corresponds usually to a minimum traffic, pre-planning mobile mission by studying the availability of satellites optimises obtained results

4.3 System calibration

The aim of this operation is to determine the parameters of transformation from scanner to IMU frame. Calibration computation is done, in general, in a static mode (Talaya et al., 2004) for example). Recognisable targets by the scanner are placed on 3 previously surveyed points. Then they are scanned by the system fixed in place.

Rotation angles from local scanner frame to the geodetic one could be calculated by one of the methods of estimating 3-D rigid transformation explained in (Lorusso et al. 1995). Only the angle about the axes Z could be verified by comparing it with the value given by Pointscape®, the software which drives the scanner (Figure 6).

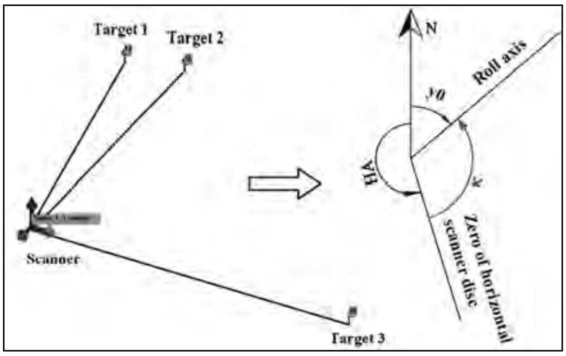


Figure 6: Simplified schema of the scanner and targets position (left), computing of one of three calibrating angles in the horizontal plane (right). $\mathbf{y_0}$: is the averaged measured yaw by the AHRS440, **HA**: calculated orientation of scanner frame. **κ**: calculated calibration value

In the same time, The AHRS is switched many times in order to obtain drift-free angles measures. It is stabilised by the GPS and the condition of no motion is integrated in the calculation

Subtracting the two triplets of angles yields directly the Euler rotation angles between scanner and IMU frames. Rotation matrix R_s^b is calculated and saved to the mathematical model. Subtraction scanner coordinates from GPS coordinates yields the precise vector from the centre of phase of antenna to the centre of sensor laser, which is another value to save with the model.

In order to verify rapidly calculated values, the mobile platform is placed in another fixed position. The scanner is operated according to its initial workflow (position determined by intersection), and then a point cloud is obtained. The georeferencing parameters are removed in order to have the point cloud in local scanner coordinate frame. Applying equation (1) results in another point cloud computed in mobile mode. The comparison between the two point clouds using ICP registration gives 2-3 cm error. It is possible to consider this value equivalent to calibration error.

5. AUGMENTATION TECHNIQUES

AHRS440 offers two augmentation methods beside GPS/INS integration. They do not involve a high cost and can change the navigation results remarkably.

- Using magnetometers: The navigation by magnetometers is usually less accurate than other techniques, but it is more stable and does not suffer from drift. So, using magnetometers is an important solution especially while GPS blockage. The main concern is not so much the surrounding environment but the mounting location on the vehicle. Most cars and trucks have a lot of steel and as a result the use of magnetometers is not possible. If one has a good mounting location, their use is possible as long as driving into warehouses or very close to large steel structures does not take place. In order to take into account the magnetic field effect of other system components, a prior standard calibration has to be done. Generally, the augmentation by magnetometers reduces random errors but could yield systematic ones. It is preferable if a verification support is available (CAD plane for example).
- Stationary yaw lock: this feature monitors GPS velocity and essentially locks the heading value when the speed falls below a given threshold. As a result, heading errors normally associated with GPS track at low speed are not seen. It is very useful in our case because the GPS grade is higher than the INS one; consequently its measures have to be taken into account with certain interest. This configuration could be considered as extension of the zero velocity update (ZUPT). The difference is employing a threshold instead of zero and depending on GPS heading rather than using the heavily accelerometer feed back to stabilize the gyro drift. Stationary yaw lock is preferable when satellites constellation is favourable.

Using both techniques together could perturb the algorithm functionality while choosing attitude correction source. Their effect could be monitored mainly by observing the change of yaw angles (about the vertical axis).

6. MOBILE POINT CLOUD

Mobile missions have to be done in go and back mode because of the one-side scan capacity of Trimble GX scanner. Walking speed has to be harmonised with that of scanner acquisition to obtain the needed resolution. In fact, these velocities affect almost only the horizontal resolution of point cloud. Vertical resolution is determined by user prior configuration allowed by Pointscape®.

TLS makes horizontal rotation during the displacement of the mobile platform. So, one can continue scanning while the incidence angle of laser beam in the object is satisfactory. We prefer usually having a small horizontal turning, related to short duration horizontal field of view, rather than large horizontal rotation of the scanner. This could be explained by the desire to obtain a point cloud with a homogenous incidence angle as shown in figure 7. The figures below show the result of scanning one part of a boulevard in Strasbourg by means of our mobile system. Figure 7a shows the raw point cloud as measured from the TLS before processing. The resulted point cloud after applying operations explained in (4.1, 4.2, and 4.3) is showed in 7b.

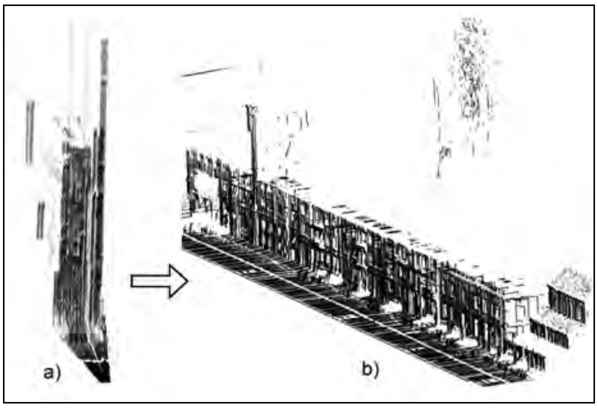


Figure 7: a: the raw point cloud in the local scanner reference frame. b: processed point cloud

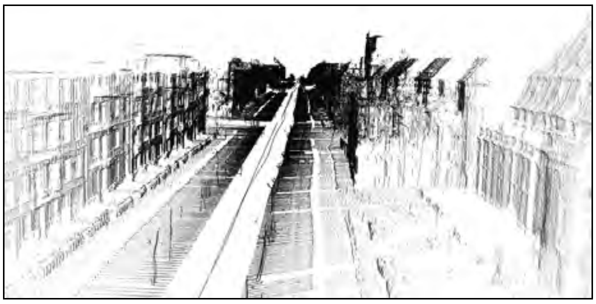


Figure 8: Perspective overall view of the scanned boulevard

6.1 Quality assessment

Mobile scanning accuracy could be quantified better by comparing it with a point cloud obtained by a static TLS or with CAD plans. The most used method is to compare some known-coordinates target points obtained by a tacheometric with mobile scanning methods.

The choice made is to compare mobile mapping point clouds with an airborne LiDAR one, regarding density and accuracy consistency. The used LiDAR point cloud has been obtained by a TopScan (Optech ALTM 1225) scanner at a density of 1.3 point/m² and an estimated accuracy of ±15cm. Mobile point cloud horizontal resolution is lower than the vertical one

because of the scanner behaviour. It is equal to 10-15 cm on the scanned object.
Overlapping Lidar and mobile mapping point clouds (figure 8) gives a good index of absolute and relative mobile scan accuracy when expressing both data in the same coordinate frame.

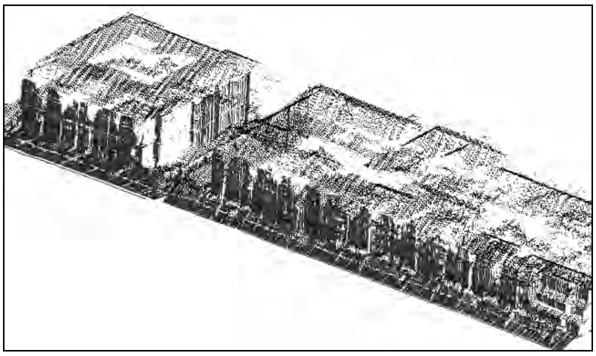


Figure 8: Overlay of airborne LiDAR (black) and mobile mapping point cloud (blue)

It could be seen clearly that the overlapping zone is very limited. Nevertheless, some cross-sections in the common zones could be drawn (gabled roof areas for example). Figure 9 shows a vertical section of 20 cm thickness in an overlapping zone. Fitting 2D straight line segments to each data set and comparing each line pair gives 32 cm as maximum distance error and about 2° as maximal direction error. Comparing many vertical profiles shows that the imprecision comes, in first place, from Z error. This could be explained by the GPS poor vertical position accuracy.

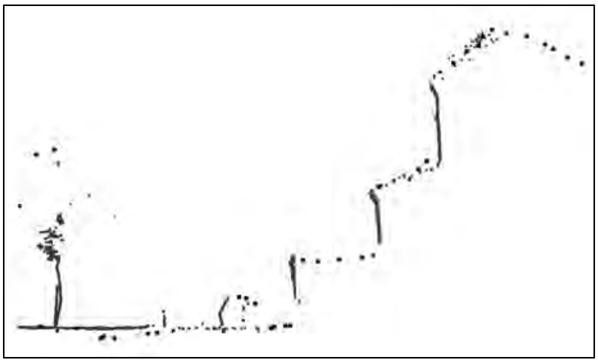


Figure 9: Vertical cross section in overlapping zone (black: airborne data, blue: mobile data)

Comparing building footprint from LiDAR and mobile scanning allows assessing planar coordinates and the effect of system heading accuracies. Linear segment comparing gives also a distance maximal error of 35 cm and a direction difference maximal of 2.9 deg. This result emphasizes on the important effect of low grade heading detectors used in the system.

6.2 Modelling process

Mobile mapping point cloud is affected by a budget of random errors from GPS, INS, TLS measurements and synchronisation error. These errors reflect directly on the relative precision of resulted point cloud. On the other hand, the resolution is not often homogeneous. It varies not only as function of the distance from the scanner to object but also of the velocity and the heading of mobile platform.
These two reasons make the modelling procedure less accurate if usual methods (as those used for static laser scanning) are applied. Hence, the interaction of user is necessary. Considering architectural knowledge to introduce geometric condition in the modeling process could be useful.
One possible enhancement of the 3D topology of mobile point cloud is to impose 3D break lines while meshing process. The constrained meshing by straight lines obliges 3D triangles to have one or two summit on this line (figure 10). This operation could product some tinny triangles around the constraining line, so one can reproduce the mesh from triangle vertexes instead of initial point cloud. An alternative solution could be obtained by a simple smoothing around inserted break lines.

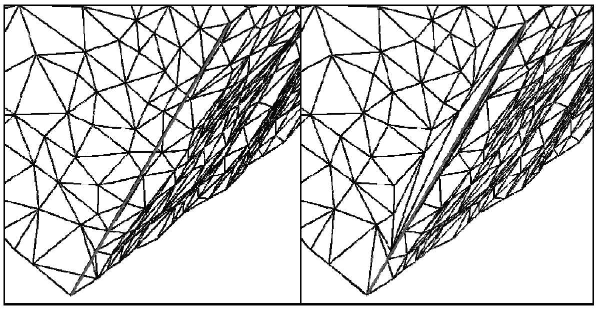


Figure 10: Introduced break line (left) and its effect on the mesh reorganisation (right)

While the point cloud is acquired from a known path, 2D meshing could yields as good result as 3D one. Points have to be projected on a surface containing the effected path and perpendicular on the scan direction. Delaunay triangulation is done between projected points but the final tessellation is done in 3D using the initial points. This method produces simpler meshing which enables more flexibility for further treatment especially from a storage size point of view.

6.3 Results

The example shown herein is a portion of a point cloud acquired at an average velocity of 0.5 m/sec. The vertical resolution was fixed at 20 cm at a distance of 100 m and the horizontal one was tenth of this value. Scanning frequency was about 500 point/sec using 4 laser shots (higher scan speed can be achieved using one shot).
The travelled path suffered from few GPS blockages equivalent to 2-3m. About 90% of GPS measured coordinates have a precision which goes below 5cm, the precision of the remaining points reaches 80 cm. Curve fitting results an estimated mean square error of ±13 cm while interpolating GPS signal absences. Magnetometers aid was used to stabilise inertial measurement, such configuration yields 1° error without external GPS aiding for 10 sec.
The final error of mobile point cloud is calculated for each point by applying the error model showed in (Alshawa et al. 2007). In present example, we state a mean error of ±20 cm. The lack of precision has to be considered in the context of mobile mapping. The total time of mobile scan is remarkably less than the one needed for a classical use of a TLS (about 20 times less).
Figure 11 shows the resulted point cloud and a constrained meshing as an example of possible modelling. The tessellation

is done in a 2D plan almost normal to scan incidence angle and containing the travelled path.

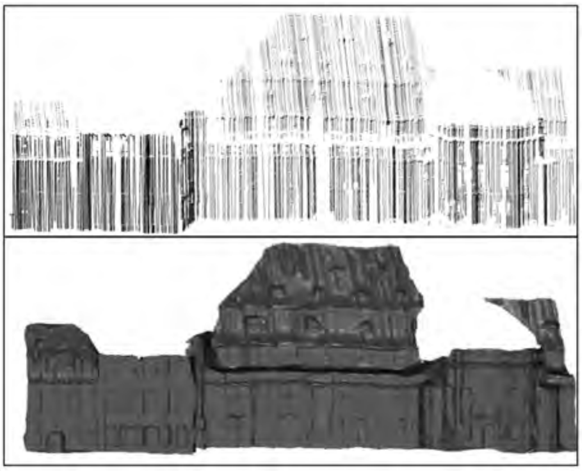


Figure 11: Possible mesh driven from a low resolution mobile mapping point cloud

7. CONCLUSION AND FUTURE RESEARCH

In this paper, a prototype of mobile mapping system based on TLS was presented. The main purpose of this work was to set out and initialize the system operation requirements and to provide some results which will be the starting point of further experiments. The reached accuracy is comparable with that of other surveying sources as aerial photogrammetry and LiDAR.
Transforming a fix terrestrial laser scanner into mobile one has proved to be feasible. The difference in accuracy between our system and the commercial ones is related to the instruments and hardware in the first place. Any enhancement (especially for inertial sensors) would be reflected directly on final point cloud accuracy.
Knowing the own navigation algorithm of the AHRS, control commands and connexion protocols of the TLS will certainly improve the operation and computing methods. While these data are generally industrial confidential, we have to suppose some hypotheses and seek for the best one.
Replacing GPS antenna by a GNSS one makes the outages shorter and enhances consequently the interpolation accuracy as well as the heading angle precision. We intend also to add a camera as another mapping sensor in order to texture the resulting models and to have some additional data which could be contribute in accuracy enhancement.
The obtained accuracy, though not comparable with the theoretical one of a fix TLS, is widely justified by the high productivity of the system. Besides, the paper has highlighted the potential improvement in the accuracy one can expect. The future work will be based on these points.

References from Journals:

Da Silva, J.F.C., De Oliveira Camargo, P. and Gallis Rodrigo, B. A., 2003. Development of a low-cost mobile mapping system: A South American experience. *Photogrammetric Record.* vol. 18, n°101, pp. 5-26

Nassar, S., Niu, X. J., El-Sheimy, N., 2007. Land-vehicle INS/GPS accurate positioning during GPS signal blockage periods. *Journal of surveying engineering* . vol:133 issues :3 Augest 2007. pp. 134 -143

Tao, C. V. Chapman, M. A. and Chaplin, B. A., Automated processing of mobile mapping image sequences. ISPRS Journal of *Photogrammetry and Remote Sensing*. vol 55, Issues 5-6, March 2001, pp. 330-346 .

References from Other Literature:

Abuhadrous, I., Nashashibi, F., Goulette, F., Laurgeau, C., Ammoun,S., 2004. Digitizing and 3d modeling of urban environments and roads using vehicle-borne laser scanner system, IEEE/RSJ Int. Conf. on Int. Robots and Systems, Sendai, Japan.

Alshawa, M., Smigiel, E., Grussenmeyer, P., Landes, T., 2007. Integration of a terrestrial lidar on a mobile mapping platform: First experiences. 5th International Symposium on Mobile Mapping Technology (MMT2007). Padova – Italy.

Hunter, G., Cox, C. and Kremer, J. 2006. Development of a commercial laser scanning mobile mapping system – StreetMapper. Second international workshop on the future of remote sensing, Antwerp.

Gajdamowicz, k., Öhman, D. and Horemuz, M., 2007. Mapping and 3D modelling of urban environment based on Lidar,GPS/IMU and image data. Proceeding of 5th International Symposium on Mobile Mapping Technology (MMT2007) Padova, Italy.

Gräfe, G., 2007. Quality management in kinematic laser scanning applications. Proceeding of 5th International Symposium on Mobile Mapping Technology (MMT2007) Padova, Italy.

Goda, C. C., 1991 The Ohio State University mobile mapping system: positioning components. The 47th annual meeting of the Institute of Navigation (ION).

Kukko, A., Andrei, C.O., Salminen, V.M., Kaartinen, H., Chen, Y., Rönnholm, P., Hyyppä, H., Hyyppä, J., Chen, R., Haggrén, H., Kosonen, I. and Capek, K., 2007. Road environment mapping system of the Finnish geodetic institute - FGI ROAMER - IAPRS Volume XXXVI, Part 3 / W52, 2007

Lorusso, A., Eggert, D.W. and Fisher, R.B., 1995. A Comparison of four algorithms for estimating 3-D rigid transformation. British machine vision conference. Vol. 1, pp. 237-246, Birmingham, England

Madeira, S., Gonçalves, J. A. and Bastos, L., 2007. Implementation of a low cost mobile mapping system. 5th International Symposium on Mobile Mapping Technology (MMT2007). Padova – Italy, 8 pages.

Talaya, J., Alamus, R., Bosc, E., Serr, A., Kornus, W. and Baron, A., 2004. Integration of a terrestrial laser scanner with GPS/IMU orientation sensors. XXth ISPRS Congress, 12-23 IAPRS Vol. 35, pp.on CD. Istanbul, Turkey, Commission 5

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

AUTOMATIC EXTRACTION OF PLANAR CLUSTERS AND THEIR CONTOURS ON BUILDING FAÇADES RECORDED BY TERRESTRIAL LASER SCANNER

H. Boulaassal*, T. Landes, P. Grussenmeyer

Photogrammetry and Geomatics Group MAP-PAGE UMR 694, Graduate School of Science and Technology
INSA, 24 Boulevard de la Victoire, 67084 STRASBOURG, France.
(hakim.boulaassal, tania.landes, pierre.grussenmeyer)@insa-strasbourg.fr

KEY WORDS: TLS, Point Cloud, Three-dimensional, LIDAR, Segmentation, Algorithm, Building, Façade, Extraction, Contours

ABSTRACT:

Since 3D city models need to be realistic not only from a bird's point of view, but also from a pedestrian's point of view, the interest in the generation of 3D façade models is increasing. This paper presents two successive algorithms for automatically segmenting building façades scanned by Terrestrial Laser Scanner (TLS) into planar clusters and extracting their contours. Since majority of façade components are planes, the topic of automatic extraction of planar features has been studied. The RANSAC algorithm has been chosen among numerous methods. It is a robust estimator frequently used to compute model parameters from a dataset containing outliers, as it occurs in TLS data. Nevertheless, the RANSAC algorithm has been improved in order to extract the most significant planar clusters describing the main features composing the building façades. Subsequently, a second algorithm has been developed for extracting the contours of these features. The innovative idea presented in this paper is the efficient way to detect the points composing the contours. Finally, in order to evaluate the performances of both algorithms, they have successively been applied on samples with different characteristics, i.e. densities, types of façades and size of architectural details. Results are satisfactory and confirm that both algorithms are reliable for forthcoming 3D modelling of building façades.

1. INTRODUCTION

There is an obvious need for creating realistic geometric models of urban areas for many application fields, such as virtual reality, digital archaeology, urban planning or GIS data bases. Therefore, the automatic reconstruction of these kinds of 3D models is of primary importance. Recently, due to its precision, reliability, degree of automation, processing speed and easy-to-handle functionalities, Terrestrial Laser Scanner (TLS) has become one of the most suitable technologies to capture 3D models of complex and irregular building façades. Indeed, based on LIDAR technology, this instrument allows recording of 3D objects in detail and produces a set of 3D points called a point cloud. Through their practicality and versatility, this kind of instruments is widely used in the field of architectural, archaeological and environmental surveying today.

Unfortunately, although techniques for the acquisition of 3D building geometries via TLS have constantly been improved, a fully automated procedure for constructing automatically reliable 3D building models is not yet in sight. This is due essentially to the difficulties of exploring directly and automatically valuable spatial information from the huge amount of 3D data. Thus many post-processing operations must be performed before accessing to reconstruction of a reliable 3D model. One of the most important operations is the segmentation. It is often prerequisite for subdividing a huge number of points into groups of points with similar properties. To deal with this subject, it is assumed in this research work, that the most prominent features of façade components are planar. The second and following operation is the extraction of contours based on these planar clusters. In this work, a contour means the set of points composing the perimeter of a planar cluster. Generally, this operation precedes the construction of the vector model.

The goal of this paper is twofold. Firstly, it aims with the automatic segmentation of TLS data into a set of planar clusters. This is achieved by applying the adaptive RANSAC (Random Sample Consensus) algorithm. Improvements are proposed here, in order to make the algorithm more efficient. Secondly, this paper presents a new algorithm for detecting and extracting planar clusters contours.

2. RELATED WORK

Over the years a vast number of segmentation methods dealing with the extraction of surfaces from laser data have been proposed. Most segmentation techniques have been developed on airborne laser data, i.e. based on 2.5D data or image data (Masaharu and Hasegawa, 2000; Geibel and Stilla, 2000), but rarely on 3D data directly. Point clouds obtained by TLS are truly 3D, especially when several scans are registered and merged. Converting such point clouds into a 2D grid would cause a great loss of spatial information (Axelsson, 1999; Gamba and Casella, 2000).

Some efficient algorithms developed initially on airborne laser data are suitable to TLS data. For instance, the works of (Pu and Vosselman, 2006; Stamos *et al*., 2006; Dold and Brenner, 2006; Lerma and Biosca, 2005) use extended region growing algorithms for extracting planar surfaces and façade features. Also (Miao and Yi-Hsing, 2004; Schnabel *et al*., 2007) propose an octree split-and-merge segmentation method to segment LIDAR data into clusters of 3D planes. Problems of techniques involving the merging operation are that the initial seed regions have a great influence on the final region. Moreover it is often difficult to decide if a region can further be extended, especially in case of noisy data. An extension of the basic region growing principle is the recover-and-select paradigm that has been introduced by (Leonardis, 95). In this approach several seed regions grow independently and result in potentially overlapping clusters. This extended approach often delivers a superior segmentation but still suffers from problems with noisy data.

In computer vision, two widely known methods are employed for shape extraction: the RANSAC paradigm (Fischler and Bolles, 1981) and the Hough transform (Hough, 1962). Both have proved that they successfully detect geometric primitives even in presence of a high proportion of outliers. However, the Hough transform is applied mainly in 2D domain, when the number of model parameters is quite small. (Tarsha-kurdi *et al.*, 2007) applied both algorithms for automatic detection of 3D building roof planes from airborne laser data. After an analytic comparison of both algorithms in terms of processing time and sensitivity to point cloud characteristics, this study shows that RANSAC algorithm is also more efficient in airborne laser data segmentation than a Hough transform.

On the other hand, the RANSAC algorithm is widely used as robust estimator of model parameters (Matas *et al.*, 2002). Moreover, the RANSAC algorithm is opposite to that conventional smoothing technique: Rather than using as much of the data as possible to obtain an initial solution and then attempting to eliminate the invalid data points, RANSAC uses as small an initial data set as feasible and enlarges this set with consistent data when possible.

Its robustness to noise and outliers renders RANSAC as a suitable choice for performing shape detection on real-world scanned data. Indeed, (Bauer *et al.*, 2005) have used RANSAC successfully to extract the main façade planes from a very dense 3D point cloud. Nevertheless, this point cloud has been obtained through image matching and was not captured by TLS. (Schnabel *et al.*, 2007) take advantages of the favourable properties of the RANSAC paradigm for detection of shapes such as planes, cylinders, spheres and torus in point clouds. Also (Tarsha-kurdi *et al.*, 2008) used successfully the RANSAC algorithm for automatic detection of building roof planes from airborne laser data. Applied on façade segmentation, (Boulaassal *et al.,* 2007) showed that a sequential application of RANSAC allows automatic segmentation and extraction of planar parts. The obtained results proved that this algorithm delivers promising results. Nevertheless, some improvements and corrections are necessary in order to make the algorithm more efficient for segmenting building façades captured by TLS.

3. DESCRIPTION OF TRIMBLE GX LASER SCANNER

Data sets used in this study have been acquired by a Trimble GX laser scanner (Figure 1). It uses time-of-flight measurement technology that is based upon the principle of sending out a laser pulse and measuring the time taken for the backscattering. Then the range distance between scanner and target is computed and combined with angle encoder measurements in order to provide the three-dimensional location of a point. Some technical specifications of this laser scanner are depicted in Table 1.

Figure 1: Trimble GX laser scanner

Technical specifications	
Distance accuracy	7 mm at 100 m
Position accuracy	12 mm at 100 m
Angular accuracy	60 µrad (Horizontal) 70 µrad (Vertical)
Grid Resolution over 360°	3 mm at 100 m with no restriction on number of points in a scan
Spot size	3 mm at 50 m
Speed	up to 5000 points per second

Table 1. Technical specifications of Trimble GX laser scanner

4. EXTRACTION OF PLANAR SURFACES USING ADAPTIVE RANSAC ALGORITHM

The adaptive RANSAC algorithm suggested by (Hartley and Zisserman, 2003) is used here in order to detect and extract planes describing planar parts of the façade. Contrary to the basic RANSAC approach introduced by (Fischler and Bolles, 1981), the adaptive RANSAC determines the number of samples adaptively. Indeed, the fraction of data consisting of outliers is often unknown. Therefore, the algorithm is initialized using a worst case estimate of outliers. This estimate can then be updated as larger consistent sets are found. Thereof, the fact of probing the data via the consensus sets is applied repeatedly in order to adaptively determine the number of samples. This operation is repeated for each sample, whenever a consensus set with a fraction of outliers lower than the current estimate is found. In this way, the number of iterations can be reduced considerably. Consequently, the improvement brought by adaptive RANSAC algorithm lies in the reduction of processing time, compared to the basic approach. Pseudo-code and more details about the adaptive RANSAC approach can be found in (Hartley and Zisserman, 2003).

5. SEGMENTATION OF FAÇADES INTO PLANAR CLUSTERS

The adaptive RANSAC algorithm is applied to extract all potential planes in form of planar clusters. As explained in (Boulaassal *et al*., 2007), the algorithm is applied sequentially and removes the inliers (valid points) from the original dataset every time one plane is detected. To determine the points belonging within some tolerance to the given plane, the Euclidian distance between each point and a plane is calculated.

In reality, data acquired by TLS are not immediately compatible with mathematical models. In other words, no planar walls, no straight edges and no right angles are directly provided in the digital model provided by the point cloud. Moreover, the raw cloud acquired by TLS has a thickness which is usually generated by noise coming from the surface roughness, the object colours, the TLS resolution capacities and the registration operation. Therefore, to detect planes representing planar façade components, a tolerance value describing the authorized thickness around a plane is imposed. The planes are thus described by planar clusters having some specific thickness.

Obviously, the quality of detected planes depends strongly on the tolerance value chosen as input. Thus, the setting of such threshold must be carefully chosen. In practice the distance threshold is usually chosen empirically as it is the case in this study. However, after (Hartley and Zisserman, 2003), this value may be computed if it is assumed that the measurement error is Gaussian with zero mean and a given standard deviation.

On the other side, the quality of planes is also related to the architectural complexity of the façade. In some cases, the RANSAC algorithm has some drawbacks which make it less efficient to obtain perfect results. Some of these problems are summarized in the next part.

5.1 Problem statement

In most cases, applying the adaptive RANSAC algorithm sequentially enables detecting all potential planes. But an important problem is encountered when these planes (walls, slabs, beams…) are intertwined (Figure 2). Indeed, when points of the best plane (the first one extracted) are withdrawn from the dataset, they cannot be affected, afterwards, to another plane. So if the first detected plane is not concerned by this problem, the others may be influenced and deformed by losing a considerable number of points already extracted by the first one detected. Indeed, affectation of those points to one plane or to another is depending on the chronological order in which the planes are detected.

To illustrate this problem, Figure 3 presents an example of intertwined planes extracted from point cloud describing the building façade shown in Figure 2. In this example, the plane depicting the wall plane (red one) was extracted firstly because it contains a large number of points. Consequently, it takes into account all points fulfilling within some threshold the flatness criterion, independently of any architectural constraint.

The points taken by the wall plane although they belong practically to other planes are drawn (in red) on the roof plane (in dark green), on the beam plane (in blue) and on the slab plane (in light green). To alleviate these defects, an algorithm aiming to optimize planes has been developed.

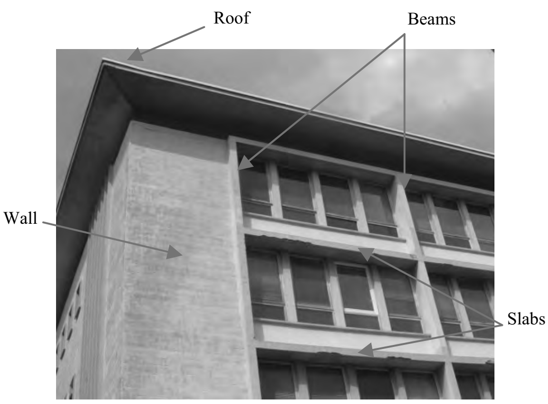


Figure 2: Example of a complex façade building composed of intertwined planes

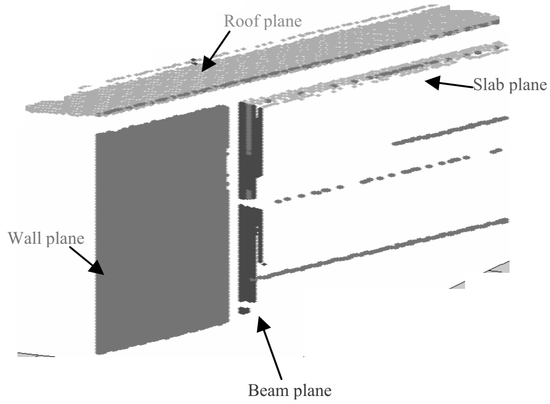


Figure 3: Example of misclassification due to chronological plane detection.

5.2 Algorithm for optimization of detected planes

In order to improve the planes resulting from the adaptive RANSAC algorithm and to make them reliable for the next step of 3D façades modeling, an optimization is required. Figure 4 describes the workflow of the developed algorithm and is detailed in following sections.

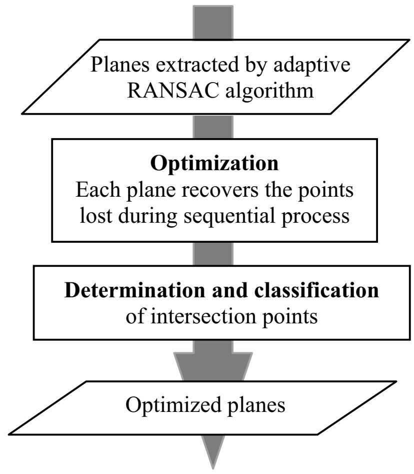


Figure 4: Workflow of the improvements brought to planes extracted by adaptive RANSAC algorithm

5.3 Processing steps for planes optimization

After extracting principal planes by the adaptive RANSAC algorithm applied sequentially, the parameters of each plane are known. Based on these parameters, the Euclidian distances between all points in the raw data and each plane are recalculated. In this way, points belonging to each plane are determined from the whole data and independently of the chronologic order in which the plane has been detected. After this improvement, previously lost points are retrieved and assigned to their corresponding plane. At this stage, it is necessary to recomputed plane parameters. Figure 5 illustrates the efficiency of the improvement. Figure 5a shows a plane before correction, i.e. with lack of points and Figure 5b after correction.

On the other hand, contrary to the planes produced sequentially by the adaptive RANSAC algorithm, the corrected ones may have common points lying for example along the intersection line between two planes. These points will be called intersection points. To be more accurate, it is necessary to detect and affect the intersection points to their right plane.

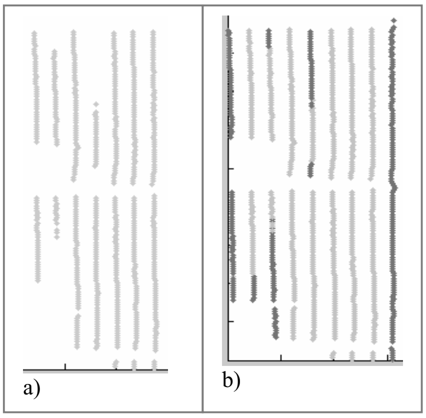


Figure 5: Detection of points belonging to a plane;
a) before correction; b) after correction (red points are retrieved)

5.4 Detection and classification of intersection points

Let's formulate a set of *N* planes by $\{PL_1; PL_2 ... PL_N\}$. Each plane PL_i with $i \geq 1 .... N$, is defined by a set of 3D coplanar points and it corresponds to a planar surface on building façade. The intersection between two planes PL_i and PL_j with $i \neq j$ exists if and only if they have at least one common point.

Once the intersection points determined, they must be affected to their right plane. To do this an algorithm of classification is developed. Let's denote $I = PL_i \cap PL_j$ with $i \neq j$ the intersection points. $C_i = (PL_i - I)$ is the complement of *I* in PL_i and $Cj = (PL_j - I)$ the complement of *I* in PL_j. The principle of this algorithm is based on the membership priority of each point according the proximity criterion. Indeed, the distance between each point of *I* and points of both complements *Ci* and C_j are calculated. Then the point is affected to the nearest one. This operation is repeated for all planes having intersection points. In this way, each point is assigned to its right plane. Therefore, the final planes are optimized. Since the number of points describing the intersection is relatively low, the time required for this processing remains negligible. Through this operation, not only the planes become suitable for the following 3D modelling, but also the definition of topological relationships between different planes is facilitated. Figure 6 shows the corrected planes corresponding to those depicted in Figure 3.

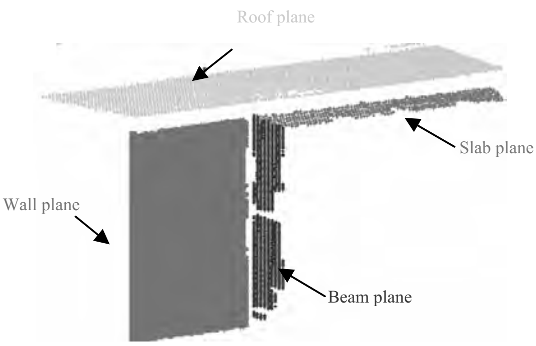
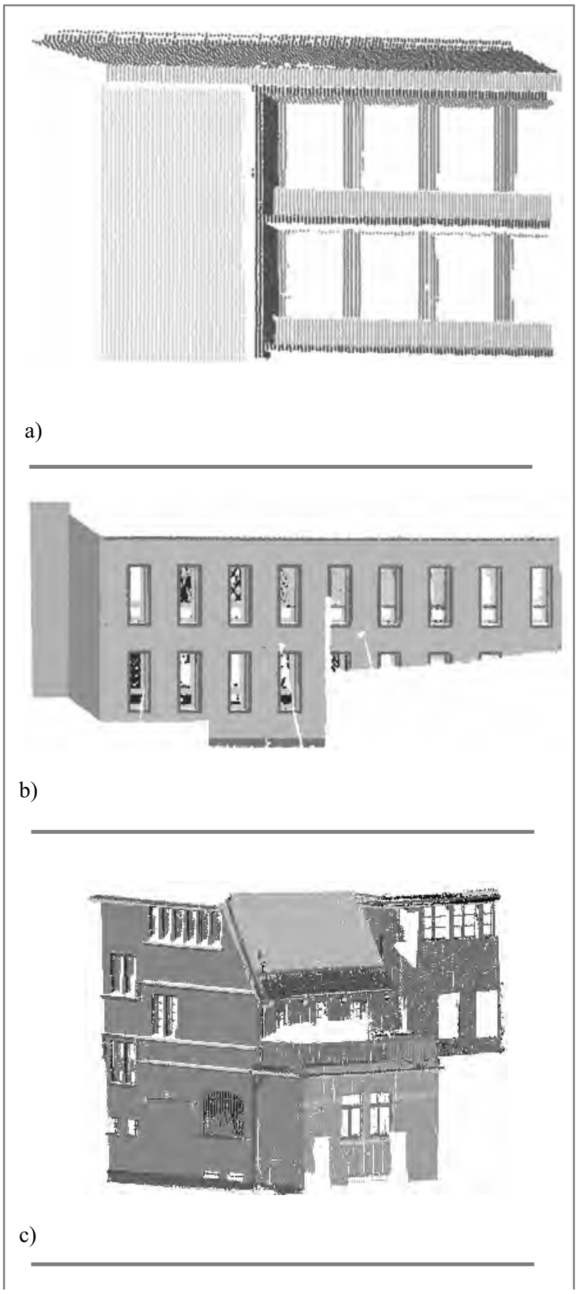


Figure 6: Effect of the improved algorithm on the planes shown in Figure 3

5.5 Application of the improved segmentation algorithm

To test the efficiency of proposed approach, the algorithm has been applied on several point clouds, captured on different types of façades. It is important to underline that the coordinates X, Y, Z are the only information used as input in this process. Façades considered in these samples are different regarding their architectural complexity as well as the type of their components (wall, balconies, beams, windows…etc.). Samples have also different characteristics regarding their density and the number of points composing each point cloud.

For instance, Figures 7 shows segmentation results of various building façades composed of features of different architectural complexities (Figures 7a, 7b, 7c and 7d). Each colour represents a plane describing a planar cluster.



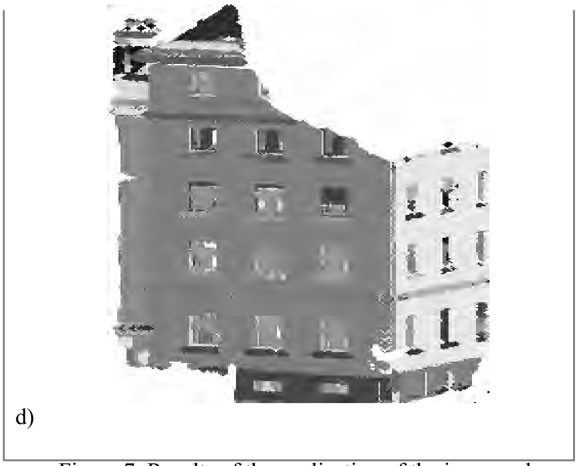


Figure 7: Results of the application of the improved segmentation algorithm on different building façades

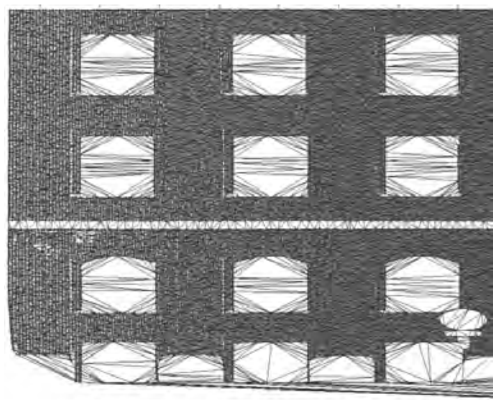


Figure 8: Delaunay triangulation of a point cloud captured on a building façade

As shown in Figure 7, the adaptive RANSAC algorithm optimized by the presented improvements provides satisfying results. The quality of segmentation may be slightly different from a data set to another depending on the characteristics of each building façade and each point cloud. Indeed, the results are better in the two first examples (Figures 7a, 7b); because the façades are entirely composed by planar surfaces and they have a simple architectural description. On the other hand, the last two samples (Figure. 7c, 7d) are composed of more than one façade. Thus, more time is needed to handle the huge amount of points. Moreover, their architectural description is more complex. Nevertheless, these results are widely sufficient for the forthcoming step, i.e. the extraction of feature contours.

6. EXTRACTION OF CONTOUR POINTS

6.1 Principle of the contours extraction algorithm

Once planar clusters are extracted by the developed segmentation approach, the extraction of their contours is carried out. To achieve this step, an efficient algorithm based on Delaunay triangulation has been developed.

Before triangulation, a new coordinate system defined in the planar cluster is determined. For this purpose, a Principal Components Analysis (PCA) is calculated based on the points of the planar cluster. The coefficients of the first two principal components define vectors that form an orthogonal basis for the plane. The third one is orthogonal to the first two, and its coefficients define the normal vector of the plane. In this principal component space, new coordinates (X_{new},Y_{new},Z_{new}) of points are calculated from the original ones (X_{origin},Y_{origin},Z_{origin}). The variance according to the third component (Z_{new}) is negligible, therefore we consider only the first two coordinates (X_{new},Y_{new}) in the Delaunay triangulation (Figure 8).

At this stage, the main new idea exploited in this algorithm is based on the hypothesis stipulating that contour points belong to the long sides of Delaunay triangles. Figure 8 shows that long sides are at the boundaries (wall contour and windows contours).The contour points are the extremities of the long sides. Therefore, lengths of all triangle sides are calculated and sorted in ascending order (Figure 9).

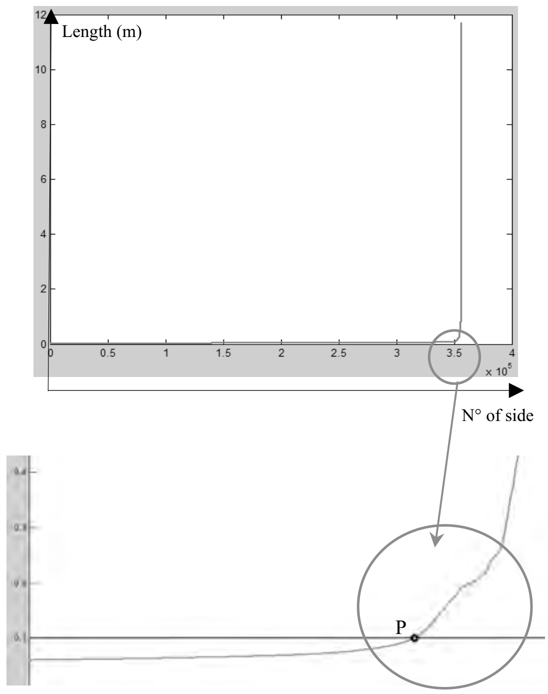


Figure 9: Curve of lengths of triangle sides resulted from Delaunay triangulation and zoom on the point fixing the threshold (threshold value set at 0.1m here)

Two classes of side lengths can be distinguished in Figure 9. The first one contains short sides that are located on the horizontal part of the curve. The second one contains long sides that are represented by the vertical part of the curve. In reality the number of points belonging to the contour should be negligible compared to the total number of points. It is confirmed by the curve in Figure 9, since the vertical part of the curve concerns only a few sides, i.e. a few triangles.

In order to implement this hypothesis it is necessary to determine the threshold separating the two classes. Typically, this value can be determined around the point P where the curvature is changing (see zoom in Figure 9). This value means that the extremities of triangle sides which length exceed 0.1m are considered as contour points.

In order to test the efficiency of proposed extraction algorithm, it has been tested over many samples. Some results are presented and discussed in following paragraph.

6.2 Application of the contours extraction algorithm

To validate the algorithm developed here, it has been applied on various planes resulting directly from the segmentation approach presented above. Planes tested are of different characteristics (point densities, total number of points, different architectural details). Figure 10 shows contour points extracted from a simple and successfully segmented plane.

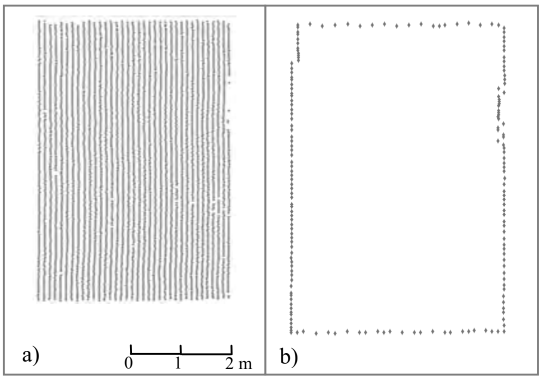


Figure 10: Example of a simple plane; a) plane obtained with segmentation algorithm; b) contour points obtained by extraction algorithm

One important advantage of this algorithm is to be able to extract automatically and simultaneously not only the outer but also the inner contours. Figure 12 and 15 illustrate this advantage through an example of plane containing outer contours (wall) and inner contours (holes like windows).

Moreover, the algorithm is able to detect the contours of architectural elements. For instance, the contours of thin elements such as transoms and mullions of windows are well extracted (Figure 12). The algorithm has also been tested on planes containing decorative architectural details (Figure 16). Despite of a high level of details like ornaments, the algorithm gives satisfactory results (Figure 17). Even in the case of destroyed façades like the walls of a medieval castle Andlau, the developed algorithm is able to detect windows. Therefore, if the quality of data and their segmentation is good, the algorithm can successfully be applied simultaneously on large and small details.

Figure 11: Plane detected by the segmentation algorithm.

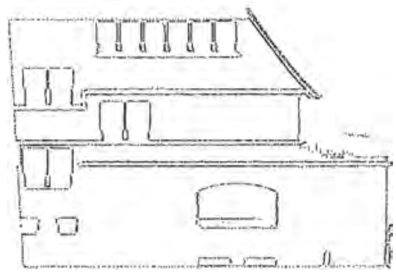


Figure 12: Outer and inner contours detected by the contours extraction algorithm applied on the plane depicted in Figure 11

Figure 13: Partial point cloud of the castle of Andlau (RGB colored)

Figure 14: Segmentation results for the point cloud shown in Figure 13

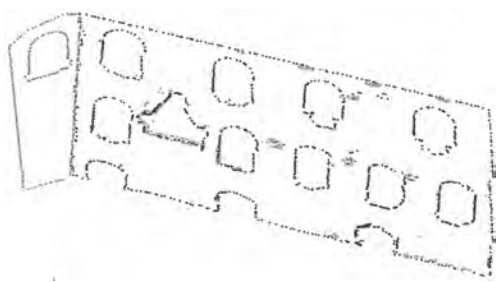


Figure 15: Outer and inner contour points detected by the contours extraction algorithm

Figure 16: Plane detection for a façade containing decorative details

Figure 17: Contour points extracted from the façade depicted in Figure 16

Figure 18 gives an example of a building composed of adjacent façades after application of the segmentation algorithm. Figure 19 shows the contour points extracted by the extraction algorithm. Two major advantages of the developed algorithms are illustrated here. Firstly, the two algorithms applied successively allow reducing the volume of points. Indeed, only points belonging to the contours are kept. Thus, the main structure of the façade is already emphasized (Figure 19).

Figure 18: Adjacent façades segmented into principal planar surfaces and features

Figure 19: Contour points of planar surfaces lying in the adjacent façades depicted in Figure 18

Secondly, these contour points simplify the work of user especially for the production of CAD models of a façade. Hence, manual or automatic digitizing is even easier in such a simplified cloud than in the whole cloud. Consequently, it allows an important gain of post processing time.

7. CONCLUSION AND FUTURE WORK

The work presented in this paper has reached the two objectives set at the beginning. Firstly it enables the automatic extraction of planar surfaces from building façades captured by TLS. Secondly, it achieves the extraction of contour points composing the boundary of each plane in order to be used afterwards in the 3D modelling. For this purpose two algorithms have been developed and presented in this paper. Their efficiency has been tested on several point clouds and several façade types. Their contribution to building façades has been discussed and analysed.

The first algorithm is derived from adaptive RANSAC algorithm and has been applied in a sequential mode in order to extract all potential planes and to handle efficiently the number of samples. Improvements were necessary, since several points are lost during the detection. Thus, corrective operations aiming to define valid points of each plane have been implemented. Several experiments have shown that the corrections were successful.

The second algorithm presented in this study aims with the extraction of the contours of planes resulting from the previous segmentation. This extraction algorithm relies on Delaunay triangulation. The innovative idea consists in analysing the length of the triangle sides in the whole triangulated network. This algorithm has been tested and applied to many samples with different characteristics. It often proved its efficiency for contour extraction**.**

The future work will exploit the results obtained by these two algorithms for the automatic reconstruction of planar features in building façades (CAD models). Therefore, it constitutes a real step for forthcoming automatic 3D modelling of building façades.

References from Journals:

Axelsson, P., 1999. Processing of Laser Scanner Data - Algorithms and Applications-, ISPRS Journal of *Photogrammetry & Remote Sensing*, vol. 54, pp. 138-147.

Hough, P., 1962. Method and means for recognizing complex patterns. In *US Patent* (1962).

Leonardis, A., Gupta, A., Bajcsy, R., 1995. Segmentation of range images as the search for geometric parametric models. *Int. J. Comput. Vision* 14, 3 (1995), 253-277.

Tarsha-kurdi, F., Landes, T., Grussenmeyer, P., 2008. Extended RANSAC algorithm for automatic detection of building roof planes from LIDAR data. *The Photogrammetric Journal of Finland*, Vol. 21, No. 1, 2008. Pp. 97-109.

References from Books:

Hartley, R., Zisserman, A., 2003. *Multiple View Geometry in Computer Vision*. pp 117-121. Cambridge University Press, second edition 2003.

References from Other Literature:

Bauer, J., Karner, K., Schindler, K., Klaus, A. and Zach, C., 2005. Segmentation of building from dense 3D point-clouds. Proceedings of the *ISPRS*. Workshop Laser scanning 2005, Enschede, the Netherlands, September 12-14, 2005.

Boulaassal, H., Landes, T,. Grussenmeyer, P,. Tarsha-kurdi, F., 2007. Automatic segmentation of building façades using terrestrial laser data. *ISPRS* Workshop on Laser Scanning 2007 and SilviLaser 2007 Espoo, September 12-14, 2007, Finland IAPRS Volume XXXVI, Part 3 / W52, 2007.

Dold, C., Brenner, C., 2006. Registration of terrestrial laser scanning data using planar patches and image data. *IAPRS*. Volume XXXVI, Part 5, Dresden 25-27 September 2006.

Fischler, M. A., and Bolles, R. C., 1981. Random Sample Consensus: A Paradigm for Model fitting with application to Image Analysis and Automated Cartography. Communications of the *ACM*, 24(6): 381-395.

Gamba, P. and V. Casella., 2000. Model Independent Object Extraction from Digital Surface Models, Proc. *International Archives of Photogrammetry and Remote Sensing, Amsterdam.*

Geibel, R. and U. Stilla., 2000. Segmentation of Laser Altimeter Data for Building Reconstruction: Different Procedures and Comparison, Proc. *International Archives of Photogrammetry and Remote Sensing, Amsterdam.*

Lerma, J.L. and Biosca, J.M., 2005. Segmentation and filtering of laser scanner data for cultural heritage. *CIPA* 2005 XX International Symposium, 26 September – 01 October, 2005, Torino, Italy.

Masaharu, H. and H. Hasegawa., 2000, Three-Dimensional City Modeling from Laser Scanner Data by Extracting Building Polygons Using Region Segmentation Method, Proc. *International Archives of Photogrammetry and Remote Sensing, Amsterdam.*

Matas, J., Chum, O., Urban, M. and Pajdla, T., 2002. RobustWide Baseline Stereo from Maximally Stable Extremal Regions. *Proceedings of the British Machine Vision Conference*, volume 1, pages 384-393, 2002.

Miao, W., Yi-Hsing, T., 2004. LIDAR data segmentation and classification based on octree structure. XXth *ISPRS* Congress, 12-23 July 2004 Istanbul, Turkey Commission 3.

Pu, S.and Vosselman, G., 2006. Automatic extraction of building features from terrestrial laser scanning. *IAPRS*, vol. 36, part 5, Dresden, Germany, September 25-27, 5 p. (on CD-ROM).

Schnabel, R., Wahl, R., Klein. R., 2007. Efficient RANSAC for Point-Cloud Shape Detection. In *Computer Graphics Forum*, Vol. 26, No. 2, Blackwell Publishing, June 2007, pages 214-226.

Stamos, I., Yu, G., Wolberg, G., Zokai, S., 2006. 3D Modeling Using Planar Segments and Mesh Elements. *3rd International Symposium on 3D Data Processing, Visualization & Transmission*, University of North Carolina, Chapel Hill, June 14-16 2006.

Tarsha-kurdi, F., Landes, T., Grussenmeyer, P., 2007. Hough-transform and extended RANSAC algorithms for automatic detection of 3D building roof planes from LIDAR data. *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*, Espoo, Finland, Sept. 12-14th. *ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information* Systems. Vol. XXXVI, Part 3 / W52, 2007.

Acknowledgments

The authors would like to thank “The Region Alsace" for supporting the MAP-PAGE UMR 694 research group and specially the topic of Terrestrial Laser Scanning.

SEMI-AUTOMATED BUILDING SURFACE TYPE EXTRACTION FROM TERRESTRIAL LASER SCANNER DATA

M. Somers [a, *], E. McGovern [a], K. Mooney[a]

[a] School of Spatial Planning, Dublin Institute of Technology, Bolton Street, Dublin, Ireland
{martin.somers,eugene.mcgovern,kevin.mooney}@dit.ie

KEY WORDS: Terrestrial laser scanning, Supervised classification, Surface type extraction.

ABSTRACT:

The identification, spatial distribution and areal quantification of building surface type are of interest in Cultural Heritage. Such information is typically gathered from on-site inspection or from photography. Both methods can be time-consuming and expensive, particularly where large numbers of buildings are involved, for example compiling a regional inventory. Frequently, such information is required in addition to geometry to more fully describe a building. The objective of this research is to develop and test algorithms for the semi-automatic extraction of building surface type from terrestrial laser scanner-acquired data.

This paper details the approach of using backscattered laser intensity with red, green, blue digital components from a terrestrial laser scanning system for a supervised classification routine. Results are presented on 3 different building façade s varying in complexity along with an overall accuracy assessment on each façade. Experimental results show that the proposed approach can be successfully executed as a semi-automatic surface extraction tool in cultural and heritage conservation.

1. INTRODUCTION

Terrestrial laser scanning (TLS) has established itself as a significant technology for recording and documenting artefacts within the realm of Cultural Heritage. Using TLS, large volumes of spatial data can be collected rapidly and with relative ease. Applications of TLS have not been confined to Cultural Heritage but have included areas such as archaeology, forensics, landform evolution and surveying.

Many commercially available TLS systems act as active sensors, collecting in addition to the geometrical information the intensity of the backscattered laser light that is used to determine the range to the object. In addition to calculating range information and recording backscattered laser intensity, commercial scanners may also record a digital colour image of the object comprising red, green and blue (RGB) channels. The intensity of the returning laser light is a function of geometrical, physical and environmental conditions. If the geometrical and environmental influences on the returned intensity can be controlled the physical influences potentially offer a considerable insight into the attributes of the surface under examination. This work proposes that classification algorithms can be used to accurately differentiate surfaces by analysing the red, green blue channels from the digital camera with the incorporation of the intensity of the returned laser signal. By leveraging the RGB information of the camera with the laser intensity it is possible to transform these datasets into bands similar to a remote sensing application. Remote sensing image analysis algorithms are particularly effective when they are based on multispectral images. A similar methodology can be extended for use with TLS data. In effect, the terrestrial laser scanner operates as a multi-spectral sensor, using the RGB and the backscatter laser radiation, allowing surface type to be extracted from terrestrial laser scanner data. A surface classification capability would add value to this maturing technology for the surveying and documenting of objects within Cultural Heritage.

Surface classification is not novel concept and is constantly under investigation with the emergence of new thematic sensors. Early work in surface categorisation e.g. Neush and Grussenmeyer (2003) centered on extracting surface features from digital images. The authors noted that a digital camera is indeed a narrow-spectral sensor and analysed the possibility of quantifying surface types using digital images.

The majority of work classifying point clouds has focused on the development of routines to extract primitive geometry elements from scans such as planes, edges and lines e.g. Belton et al., (2006). Classification routines (Lichti, 2005) within point clouds have been demonstrated in the differentiation of object/surface types. This research outlined how thematic class labels were assigned to data points in the point cloud in essence assigning an informational type to the point. Laser scanning systems are not restricted to using a digital camera as a multispectral tool, multispectral laser scanning systems are currently being developed (Hemmleb et al., 2006). This particular system contains four independent laser modules that are used to detect damage on building surfaces by using a supervised classification routine to ascertain regions of variable moisture content.

Research in analysing backscatter laser light has been extensive in the preceding years with a number of groups from various backgrounds contributing to the field, for example Ellis et al., (2002) and Ruiz-Cortes & Dainty (2002). This has led to new approaches of extracting information from the backscattered intensity, for example in remote sensing derived correction models for the backscattered intensity e.g. Jutzi et al., (2003), Thiel & Wehr, (2004) and Hofle & Pfeifer, (2007). Previous research has mainly focused on airborne laser scanning but similar principles could also be applied to TLS in deriving mathematical models that approximate the amount of backscatter laser intensity returned to a scanner. Most of the research work with TLS has focused on geometric accuracy e.g. Lichti, (2007) and Reshetyuk (2006). Exploratory work in ascertaining the factors that effect backscattered intensity in TLS have taken place (Pfeifer et al., 2007). In their work the authors used Lambertian targets to relate the returned laser signal to factors such as colour and range with the intention of quantifying the intensity within a TLS system.

For Cultural Heritage, building type information extracted from terrestrial laser scan can be an invaluable tool for inventory documentation of buildings. To establish such a method it is important to understand the spectral responses that are taking place within the scanner. The red, green and blue channel from the camera relates the absorption and reflection of light in the visible region of the spectrum, the backscattered laser intensity holds inherit information on the reflectance of the material based on the surface finish and colour (Hanke et al., 2006). When determining categorisation of the pixels for the informational classes each pixel is assessed via a temporal relationship with pixels in surrounding bands. Image classification is the process of applying informational elements or descriptive terms that are more easily recognisable from multi-spectral data. Primarily used in remote sensing (Swain, 1978), image classification uses statistical analysis to identify natural groups or structures with similar attributes and characteristics. For the purpose of this research supervised classification strategies are applied to point cloud datasets and rely on an analyser/interpreter to lead the classifier in identifying areas of the image that are known to belong to an information class.

2. METHODOLOGY

This paper applies a semi-automated surface classification algorithm workflow from the field of remote sensing to TLS systems to determine building surface composition and spatial variability. All point cloud datasets were acquired using a Trimble GS200 system, the specifications for which are given in **Table 1**. Three buildings were selected to test the feasibility of extracting surface information using supervised classification. Building A contained two informational classes (red brick and grout). Building B contained four informational classes (red brick. brown brick, grout and granite) and Building C contained four informational classes (red brick, a painted surface (cream colour) and two different variations of sandstone). Thus the complexity in the number of informational classes increased from building A through C.

To assess the successfulness of using a commercial terrestrial laser scanning system directly for building classification the laser image (backscatter intensity) is used to classify a building type and compared to the merged data set of the laser image and the digital camera (RGBI) of the scanner. Building façades were used to construct datasets, these data sets were single objects therefore no targets were required for point-cloud registration. All scans were recorded at a nominal 5 mm spatial resolution on the object with the scanner approximately 10-20m from the scanning surface.

Standard Range	2-200m
Scanning Speed	5000pts/sec
Minimum Resolution	3mm @ 100m
Laser Beam Diameter	3mm @50m
Field of View	360 H, 60 V
Camera Resolution	768*576 colour resolution

Table 1. Trimble GS200 specifications.

After scanning the point clouds were processed using Trimble Realworks 6.2 software. Regions of interest were segmented from the parent clouds and the data exported via ascii format. Seven parameters per point were exported from the cloud. These were the x, y, z coordinates, the backscattered laser intensity and the red, green and blue channels from the digital camera. For the purpose of simplicity image registration of the laser image and digital images from the camera was not taken into account, the RGB values from the point clouds are assumed to match their respective pixels of the laser image. These exported ascii files were then further manipulated via Matlab (Version 7.04) software to generate four separate images (often referred to as bands in remote sensing). These four images were raster images of the red, green, blue channels from the camera and the scattered laser intensity, respectively, and formed the basis for the classification routine. Once the four images/bands had been created they were imported into ERDAS Imagine software which is a standard program used in remote sensing and comprises a range of tools for processing satellite imagery. For the dataset to be of use within ERDAS Imagine each of the 4 images/bands are converted to a surface object. Surface object in ERDAS pertain to planar datasets. Typically ERDAS Imagine surface objects are derived from satellite datasets where each element of the surface object represents the x, y spatial values of the scan. As the TLS system is operating in a front-on view, the x, z values are substituted for x, y, z values. The newly created surfaces were saved in .img format. This is a format native to ERDAS Imagine and classification is carried out on these .img files. To complete the pre-processing stage, individual images, representing the separate spectral bands, were stacked to create the multispectral .img files for each building.

As previously mentioned classification is the process of dividing pixels into a finite number of informational classes or categories based on their mutual characteristics. For supervised classification to function the information class should be known prior to initialising the classification routine. Therefore a signature of the available information classes must be created. Five regions within the image were used for each informational class in composing the signature. This process of grouping five regions of similar type is repeated for all remaining informational classes that are available in the image. Once the signature file has been built the supervised classification algorithm is initiated. A maximum likelihood classification algorithm was used for the supervised classification routine.

Accuracy assessment was carried out to validate that the proposed information classes from the classification algorithm have been assigned correctly. The accuracy assessment is based on the fact the user is aware of the true classification value and can validate the fact that the classification is indeed correct or the classification has not been established correctly. Typically the known class is referred to as a reference class. Reference classes were gathered from a reference images that were taken at the same time as the scan. A series of random points were used to derive if the informational classes were indeed correct. By using points selected at random bias in the calculated accuracy is eliminated. The distribution parameters for the random points used were stratified random distribution. The number of pixels that is used in the calculated is important. It has been established (Congalton, 1991) that a minimum of 250 pixels per informational class are required to estimate the mean accuracy to within plus or minus 5 %. Once the distribution of points had been assigned accuracy assessment of the classes was carried out.

3. RESULTS

Three different building surface types were selected to assess the viability of using the red, green and blue channels of the commercial TLS camera (Trimble GS-200) with the backscattered laser intensity, RGBI, these sites shall be referred to as Building A, B and C. These RGBI images are compared to a single band, namely the backscattered intensity band for comparison.

3.1 Building A

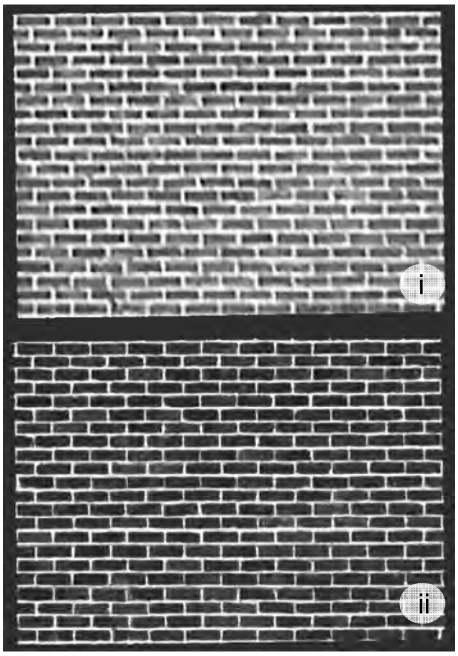


Figure 1. Building A - RGB camera (i) laser intensity (ii).

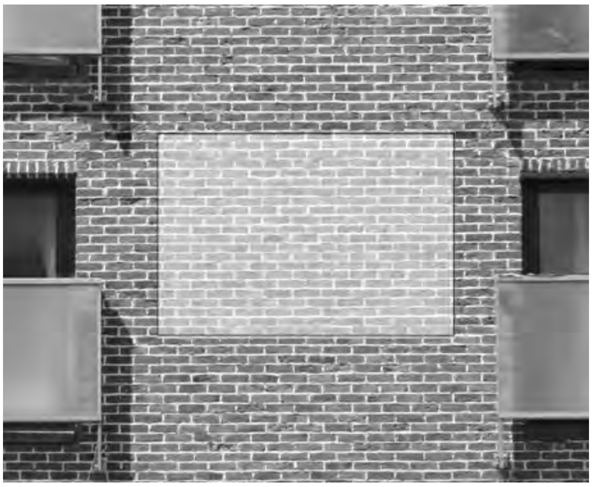


Figure 2. Building “A” scan area.

The façade of Building A comprised of two information classes red brick and grout (Figure 2). There are slight pigment variations within the bricks, these variations were assumed to be insignificant in the classification of the brick informational class of building A.

The results of the classification process are shown in Figure 3, Brick-red (red) and Grout (yellow). The classified image Figure 3(i) is derived from the RGBI stack, the classified image Figure 3(ii) is calculated by using the backscattered intensity recording. A series of point locations on the reference image were selected in a stratified random configuration to construct an error matrix and calculate the accuracy assessment. **Table 2**, displays the error matrix using the informational classes outlined previously, classified with RGB and intensity (RGBI).

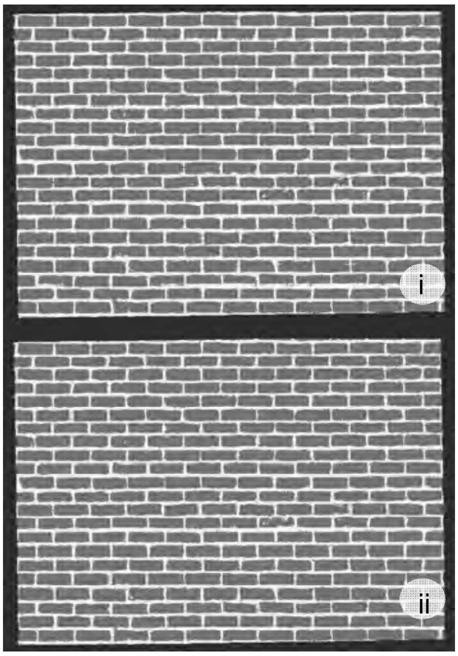


Figure 3. Classification – RGBI (i), laser (ii)

Error Matrix

Classified Data	Unclassified	brick-red	grout	Row Total
Unclassified	24	0	0	24
brick-red	4	616	72	692
grout	0	16	292	308
Column Total	**28**	**632**	**364**	**1024**

Accuracy Assessment

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	28	24	24	---	---
brick-red	632	692	616	97.47%	89.02%
grout	364	308	292	80.22%	94.81%
Totals	**1024**	**1024**	**932**		
Overall Classification Accuracy =	91.02%				

Kappa Statistics
Overall = 0.8110, Unclassified = 1, brick-red = 0.7131, grout = 0.9194.

Table 2. Building A - classification using RGBI.

Data pertaining to the two informational classes of Building A with supervised classification using backscattered laser intensity, **Table 3**.

Error Matrix

Classified Data	Unclassified	grout	red-brick	Row Total
Unclassified	32	0	0	32
brick-red	0	56	668	724
grout	0	228	40	268
Column Total	**32**	**284**	**708**	**1024**

Accuracy Assessment

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	32	32	32	---	---
brick-red	708	724	668	94.35%	92.27%
grout	284	268	228	80.28%	85.07%
Totals	**1024**	**1024**	**928**		
Overall Classification Accuracy =	90.63%				

Kappa Statistics
Overall = 0.7858, Unclassified = 1, brick-red = 0.7494, grout = 0.7935.

Table 3. Building A - classification with laser intensity.

3.2 Building B

Figure 4. Building "B" scan area.

Building B contained four information types grout, granite and brick that were further divided into subsets containing colour variations of red and brown. It must be noted that the class containing brown brick, also contained grey coloured bricks that are randomly distributed within the façade (**Figure 4**). The resulting images prior to classification within ERDAS imagine are shown in **Figure 5**.

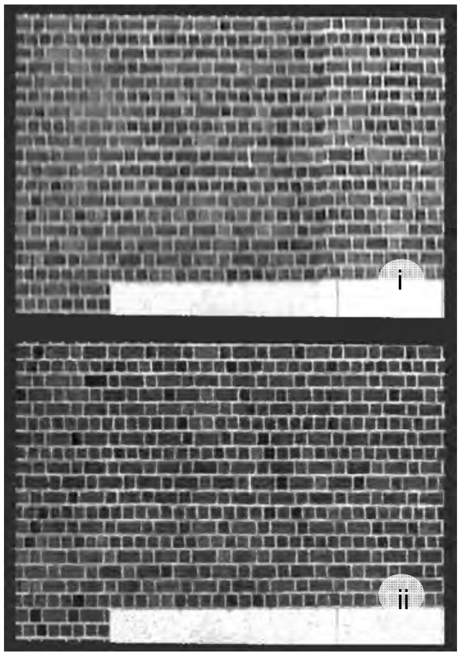


Figure 5. Building B - RGB camera (i), laser intensity (ii)

The top classified image of **Figure 6**(i) represents the RGB-intensity stack, the image below **Figure 6**(ii) classification derived from the backscattered intensity recording.

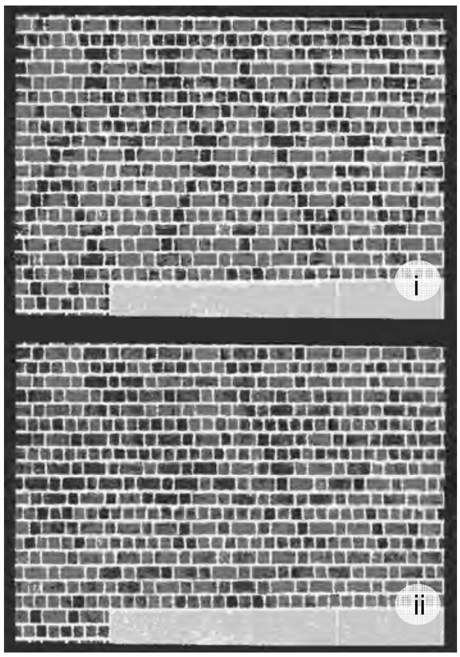


Figure 6. Classification – RGBI (i), laser (ii).

The error matrix and the accuracy report post classification on the RGBI image **Figure 6**(i) is displayed in **Table 4**. Each of the informational class is represented by a colour as follows, Brick-red (red), Brick-brown (blue), Grout (yellow) and Granite (green).

Error Matrix

Classified Data	Unclassed	Granite	Grout	Br-Red	Br-Brn	Row Total
Unclassified	7	0	0	0	0	7
Granite	0	265	0	2	0	267
Grout	1	4	312	20	5	342
Brick - red	0	1	32	303	37	373
Brick - brown	0	0	49	29	213	291
Column Total	**8**	**270**	**393**	**354**	**255**	**1280**

Accuracy Assessment

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	8	7	7	---	---
Granite	270	267	265	98.15%	99.25%
Grout	372	373	303	81.45%	81.23%
Brick - red	375	342	312	83.20%	91.23%
Brick - brown	255	291	213	83.53%	73.20%
Totals	**1280**	**1280**	**1100**		

Overall Classification Accuracy = 85.94%

Kappa Statistics

Overall = 0.8119, Unclassified = 1, Granite = 0.9905, Grout = 0.8759, Brick-red = 0.7354, Brick-brown = 0.6653.

Table 4. Building B - classification using RGBI.

When looking at just the backscattered intensity on the same four informational classes from **Figure 6**(ii), the corresponding error matrix and accuracy assessment are shown in **Table 5**.

Error Matrix

Classified Data	Unclassed	Granite	Grout	Br-Red	Br-Brn	Row Total
Unclassified	6	0	0	0	0	6
Granite	0	262	3	1	0	266
Grout	0	0	303	15	34	379
Brick - red	0	6	54	291	10	334
Brick - brown	0	1	31	87	176	295
Column Total	**6**	**269**	**391**	**394**	**220**	**1280**

Accuracy Assessment

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	6	6	6	---	---
Granite	269	266	262	97.40%	98.50%
Grout	391	334	303	77.49%	90.72%
Brick - red	394	379	291	73.86%	76.78%
Brick - brown	220	295	176	80.00%	59.66%
Totals	**1280**	**1280**	**1038**		

Overall Classification Accuracy = 81.09%

Kappa Statistics

Overall = 0.7465, Unclassified = 1, Granite = 0.981, Grout =0.8664, Brick-red = 0.6646, Brick-brown = 0.5129.

Table 5. Building B - classification with laser intensity.

3.3 Building C

The third and final building selected increases the complexity available with building A and building B. There are a number of different surface types available within the field of view of the scanner. Red brick, sandstone and paint. Two different types of sandstone are visible in the scan (**Figure 7**). The paint on the façade is cream in colour and has similar tonal appearance to the surrounding sandstone.

Figure 7. Building C - RGB image – external camera.

Figure 8. Point Cloud, Building C – intensity (left), RGB (right).

Figure 8 shows the backscatter intensity on the left and the RGB image on the right, notice slight colour changes occurs across the RGB image a magnified image is shown in **Figure 9** and its backscattered intensity **Figure 10**.

Figure 9. Building C - RGB image.

Figure 10. Backscattered Intensity represented in greyscale prior to classification.

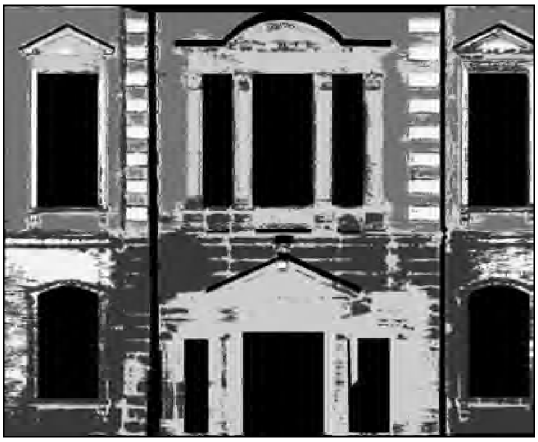


Figure 11. Building C - Classification with RGBI.

Figure 12. Building C - Classification with intensity.

One aspect of note from building C is the difference in spatial resolution in the scanner image (**Figure 9**) compared to backscattered intensity the greyscale image of **Figure 10**. The fine detail available from the intensity image isn't available in the RGB image (Notice the grout lines on the brick work are not visible). The is due to the fact that the setting in scanner where not optimised to use the highest projected images that were available, image taken at 100% zoom setting.

The RGBI classified image of Building C is shown in **Figure 11**, Brick-red (red), Paint (yellow), Sand-I (green), Sand-II (blue). Its corresponding error matrix and accuracy report are available in **Table 6**.

Error Matrix

Classified Data	Unclassed	Br-Red	Paint	Sand-I	Sand-II	Row Total
Unclassified	166	0	0	0	0	166
Br-Red	1	318	1	2	0	322
Paint	0	0	299	0	0	299
Sand-I	0	50	30	311	8	399
Sand-II	0	10	204	54	82	350
Column Total	167	378	534	367	90	1536

Accuracy Assessment

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	167	166	166	---	---
Br-Red	378	322	318	84.13%	98.76%
Paint	534	299	299	55.99%	100.00%
Sand-I	367	399	311	84.74%	77.94%
Sand-II	90	350	82	91.11%	23.43%
Totals	**1536**	**1536**	**1176**		

Overall Classification Accuracy = 76.56%

Kappa Statistics

Overall = 0.7074, Unclassified = 1, Brick-Red = 0.9835, Paint = 1, Sandstone I = 0.7102, Sandstone II = 0.1866.

Table 6. Supervised classification with RGBI Building C.

The backscatter intensity from Building C is shown in **Figure 12** and its corresponding error matrix and accuracy report in **Table 7**. Brick-red (red), Paint (yellow), Sand-I (green), Sand-II (blue)

Error Matrix

Classified Data	Unclassed	Br-Red	Paint	Sand-I	Sand-II	Row Total
Unclassified	82	0	0	0	0	82
Br-Red	0	281	4	14	2	301
Paint	0	0	290	0	0	290
Sand-I	0	7	25	261	22	315
Sand-II	0	4	118	50	118	290
Column Total	82	292	437	325	142	1278

Accuracy Assessment

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	82	82	82	---	---
Br-Red	292	301	281	96.23%	93.36%
Paint	437	290	290	66.36%	100.00%
Sand-I	325	315	261	80.31%	82.86%
Sand-II	142	290	118	83.10%	40.69%
Totals	**1278**	**1278**	**1032**		

Overall Classification Accuracy = 80.75%

Kappa Statistics

Overall = 0.7521, Unclassified = 1, Brick-Red = 0.9139, Paint = 1, Sandstone I = 0.7701, Sandstone II = 0.3328.

Table 7. Supervised classification – intensity, Building C

4. ANALYSIS

The accuracy report contains the producer's and user's accuracy. Producer's accuracy refers to the percentage of a surface type that has been correctly classified; user's accuracy refers to the percentage of a classified type that is truly of that type. The producer's accuracy represents how accurate the original surface has been classified, the user's accuracy define how good the end classification has been, i.e. the classified map. Also presented is the derived value of kappa. The kappa value represents the proportion of the measurand obtained after removing the proportion of agreement that could be expected to occur randomly; typically values lie in the range between 0 and 1. (0 - 0.2) slight agreement, (0.2 - 0.4) fair agreement, (0.4 - 0.6) moderate agreement, (0.6 - 0.8) substantial agreement, (0.8 – 1) almost perfect agreement (Landis & Koch, 1977).

4.1 Building A

Comparing both images of **Figure 1**, the laser intensity image is superior in resolution compared to the RGB from the camera in the TLS system. For the most part there is increased benefit of using an RGBI data set compared to using just a single backscattered intensity band. Overall Kappa values for RGBI and intensity are 0.8110 and 0.7858. The producer's accuracy is very similar in both cases classification using RGBI (brick-97% grout-80%) and classification using intensity – (brick-94% grout-80%).

Of note are the differences in the kappa value of RGBI and the backscatter intensity of grout 0.9194 to 0.7935. The deviation in values would suggest that the RGBI image is better in differentiating grout compared to using just the laser intensity image. On a whole RGBI statistical classification values are marginally higher than intensity based values.

4.2 Building B

For the classification of building B four informational classes were proposed, granite, grout, red brick and brown brick though the brown brick information class contained elements that were somewhat grey in colour (**Figure 4**). Therefore the assignment of "Brick-brown" to the informational class is solely used to differentiate it from the "Brick-Red" class. Compared to Building A there is an increase in complexity and deviation in the brick composition that is easily recognisable as a diamond pattern on the façade. Again comparing to Building A the brick class is distinct from the grout though with smaller values in kappa for the RGBI, grout kappa value has decreased slightly to 0.8759.

Using the RGB image from the internal camera of the scanner (**Figure 5**i) it is seen that the granite class (lower right of image) and the grout class are visibly similar, though there are significant differences when classified with the classification routine, for the RGBI scan, of 267 classified granite pixels none were assigned a grout class. This can be extended by looking at the grout class where four pixels from 342 pixels were misclassified as granite. This leads to high values of user's accuracy for granite (99.25%) and grout (81.23%). This lower value for grout can be attributed to 25 pixels from both brick classes being perceived as being grout.

The red brick class suffers from a lower kappa compared to Building A, reasoning that there is misclassifications taking place with the increased number of informational classes. For example, classification based on RGBI (**Table 5**) out of a total 373 points that were classified as "Brick-red" 303 were correctly classified with the remaining pixels falling equalling into grout-32 and brown brick-37. When taking the intensity into account out of 334 points, 291 points were classed correctly with only 10 being classed as brown brick and 54 being classified as grout. This is visually shown in **Figure 6**. Where the alternating diamond pattern is emerging from the red background compared to just using the backscattered intensity **Figure 6** (ii) where the pattern is partially visible. From the table it can be concluded that the RGBI image correctly classifies more informational class than just using the laser intensity.

4.3 Building C

Building C represents a more complex example compared to building A and B. There are a number of unclassified points in building C, this is due to one of two reasons, segmentation of the point cloud i.e. removal of unwanted points and object from the image (windows and a lamp post). Also laser shadow is present in the images, meaning some building features are occluded from the scanner.

From the error reports the informational classes derived using intensity performed better for all classes but for the red brick informational class. For RGBI we get accuracy values of correctly classified points of red brick (99%), paint (100%) sand stone I (78%) sand stone II (23%) comparing to backscattered intensity red brick (93%), paint (100%) sand stone I (82%) sand stone II (40%). In both cases the informational class of paint is perfect (100% and 100%) compared to the accuracy of the other classes. This is due to the fact that all paint pixels were indeed paint pixel. Though looking at figure 11 and figure 12 regions that are indeed paint have been misclassified hence why the producer's accuracy is significantly lower 56% and 66% respectively.

From figure 11 we see that three of the informational classes (red-brick, sand-stone I and sand-stone II) had been classified correctly (high values in user and producer's accuracy), though there are variations in the classification around the window areas in the upper part of the screen, this is due to localise shadows darkening the region and misclassifying pixels. Overall the classification is correct and is reflected in the high values obtain of producers accuracy of Table 12.

The RGBI classification performs poorer to the results obtained for Building A and Building B compared to the backscattered intensity. From Building C it is evident that lighting conditions play a significant role in the how successful the classification routine can be, for instance the prominent area of misclassification occurs around the window area where shadowing effects the performance of the algorithm.

5. CONCLUSION AND OUTLOOK

A supervised classification of building surfaces was carried out on real-world terrestrial laser scanning data. From these datasets a material type classification map was extracted and areas of different material type located. It has been successfully shown that it is possible to acquire semi-automated surface classification from TLS point cloud data using remote sensing classification algorithms, with accuracy rates of up to 97% when using RGBI for two information classes (Building A). Using this approach of a semi-automatic means of extracting a building surface type could be used in a range of applications, predominantly in area/volume extraction and site inventory analysis.

One aspect of improving this method for building classification is the need to overlay images that are comparative in resolution. From the images provided, the difference in resolution of the native RGB image to the laser intensity is noted. While it is possible to capture images at greater resolution with the scanner some image matching is required. At the moment the resolution of the laser imaging is superior to the RGB image outputted by the camera available with the scanner. Therefore further work shall focus on using higher resolution images from the current TLS system and or an external digital camera.

Resolution aside the method of stacking RGB image with the laser image does show that it is possible to classify pixels into informational type. The scan data from Building B highlights the importance of using the RGBI instead of just the backscattered intensity image in that colour variations within individual bricks may be classified to a better degree by using the RGBI merged image to the backscattered laser image.

6. REFERENCES

Congalton, R. (1991) A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of Environment 37: 35-46.

Belton, D. & Lichti D. (2006) Classification and segmentation of terrestrial laser scanner point clouds using local variance information, ISPRS Commission V Symposium 'Image Engineering and Vision Metrology', Dresden, Germany, 2006, pp. 44-46

Ellis, J., Caillard, P. & Dogariu, A. (2002) Off-diagonal Mueller matrix elements in backscattering from highly diffusive media. Journal of the Optical Society of America A, 19, 43-48.

Hanke, K., Grussenmeyer, P., Grimm-Pitzinger, A. & Weinold, T. (2006) First experiences with the Trimble GX scanner.

Hemmleb, M., Weritz, F., Schiemenz, A. & Grote, A. (2006) Multi-spectral data acquisition and processing techniques for damage detection on building surfaces. , ISPRS Commission V Symposium 'Image Engineering and Vision Metrology', Dresden, Germany, 2006.

Hofle, B. & Pfeifer, N. (2007) Correction of laser scanning intensity data: Data and model-driven approaches. ISPRS Journal of Photogrammetry and Remote Sensing, 62, 415-433.

Jutzi, B., Eberle, B. & Stilla, U. (2003) Estimation and measurement of backscattered signals from pulsed laser radar: Proc. SPIE, 4885, 256-267.

Kersten, T., Sternberg, H. & Stiemer, E. (2005) First Experiences with Terrestrial Laser Scanning for Indoor Cultural Heritage Applications Using Two Different Scanning Systems. Proceeding of the ISPRS working group V/5 'Panoramic Photgrammetry Workshop'.

Landis, J. R. & Koch, G. G. (1977), The measurement of observer agreement for categorical data, Biometrics. Vol. 33, pp. 159--174

Licti, D. (2005) Spectral Filtering and Classification of Terrestrial Laser Scanner Point Clouds. The Photogrammetric Record, 20, 218-240.

Lichti, D. (2007) Error modeling, calibration and analysis of an AM-CW terrestrial laser scanner system. ISPRS Journal of Photogrammetry and Remote Sensing, 61, 307-324.

Neush, T. & Grussenmeyer, P. (2003) Remote sensing object-orientation images analysis applied to half-timbered houses. CIPA XIXth International Symposium.

Pfeifer, Dorninger, Haring, Fan, (2007), Investigating terrestrial laser scanning intensity data: quality and functional relations. Proc. of VIII conference on Optical 3D measurement tech.

Reshetyuk, Y. (2006) Calibration of terrestrial laser scanners for the purposes of geodetic engineering. 3rd IAG Symposium on Geodesy for Geotechnical and Structural Engineering and 12th FIG Symposium on Deformation measurement.

Ruiz-Cortes, V. A. & Dainty, J. C. (2002) Experimental light-scattering measurements from large-scale composite randomly rough surfaces. Journal of the Optical Society of America A, 19, 2043-2052.

Swain, P. H. & Davis, S. M. (1978) Remote Sensing; The quantative Approach. McGraw-Hill Inc.

Thiel, K. H. & Wehr, A. (2004) Performance capabilities of laser scanners--an overview and measurement principle analysis. International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences, 36, 14-18.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

MARKLESS REGISTRATION FOR SCANS OF FREE-FORM OBJECTS

A. Valanis [a], Ch. Ioannidis [b]

Laboratory of Photogrammetry, National Technical University of Athens, Greece
[a] artvalanis@yahoo.gr
[b] cioannid@survey.ntua.gr

KEY WORDS: Registration, ICP initialization, Cultural heritage, Free-form objects, Constructions

ABSTRACT:

Rather frequently users are faced with the challenge of how to register scans of free-form objects in the absence of targets, especially in cultural heritage and industrial applications. In such cases, the surface of the objects of interest is usually smooth, with no characteristic points (e.g. sculptures, industrial constructions etc.) and the acquisition process is often subdued to space limitations, like small object-instrument distance, object height greater than object-instrument distance, difficulty in placing targets above a given height etc. Moreover, unless the scans to be registered are already approximately aligned, most commercial software cannot initialize the ICP process without at least three common targets or conjugate points. The proposed approach enables the initialization of ICP for scans with no common targets, when conjugate points are hard to identify. Specifically, a repetitive process estimates the distance between two point-clouds as one of those is continuously rotated according to a sampling strategy. When a sign-change of the estimated point-cloud distance is detected, the search space is reduced and the approximate unknown angle is derived. A thorough validation of the proposed process is presented here along with an example application for the complex of sculptures of the monument of Zalongon.

1. INTRODUCTION

Terrestrial laser scanners are increasingly being used for various types of applications. This type of equipment has undeniably enabled the fast and detailed capture of the 3D form of various objects of interest. Currently, several systems with varying fields of view provide users with various levels of resolution and accuracy, and the ability to capture intensity and/or color. Normally, in order to model an object, several scans are acquired and registered with the use of 2D or 3D targets.

There are many cases e.g. cultural heritage and industrial applications, where the acquisition process is subdued to space limitations i.e. small object-instrument distance, object height greater than object-instrument distance, difficulty in target placing above a given height etc. Consequently, data acquisition most often involves a few scans that contain no targets but need to be integrated into processing so that the final model be complete. In the absence of targets, conjugate points are selected and the Iterative Closest Point (ICP) algorithm is applied so as to align the point clouds (Besl and McKay, 1992). That practice is difficult or even impossible to apply for cases of free-form objects whose surface is smooth or even uniform and no characteristic points of interest can be identified.

This paper suggests a solution in order to bypass this problem and actually initialize the ICP process without the use of targets or conjugate points. An early version of the proposed approach was designed in order enable mark-less registration of point-clouds for a cooling tower of the Public Power Corporation of Greece but it has never been presented in detail. Recently, a project that involved the survey of the Zalongon monument (a 17.5x 12.5m complex of sculptures) called for the application of such an approach. Based on the experience gained from the cooling tower project, the acquisition process was especially adjusted so as to enable the application of a similar solution and refine the original algorithms so as to develop a methodology applicable for similar problems. Therefore, beside the data that were required for the survey of the monument, data that would be used in order to test and validate the proposed approach were also acquired. In this contribution the authors describe the proposed methodology in detail, present the experimental results derived for the validation of the proposed approach and discuss the technical aspects and the results of the application for the case of the Zalongon monument are presented in detail.

2. INITIALIZATION OF ICP PROCESS

ICP compares two point clouds and continuously transforms one of them until the process converges, i.e. until the objective function calculated reaches a minimum. In order for the ICP to provide a solution either the two scans to be registered have to be in approximate alignment, or the approximate transformation between the two scans must be derived i.e. calculated from the equations formed for at least three conjugate points selected by the user. Therefore, if no targets can be used, either dense scans (i.e. scans acquired from nearby set-ups) need to be acquired or only the required scans may be acquired provided that characteristic points can be recognized between any two given scans.

These two solutions are the most popular among the users of commercial software such as Geomagic™ and Cyclone™, which enable point-cloud registration by means of the ICP. In both of these programs the user may perform registration without the use of targets either based on scan world coordinates for the case of dense scans or on the selection of conjugate points in the case of sparse scans. However, both of the aforementioned solutions present several serious drawbacks and very often lead to pour or no results. Specifically, by acquiring dense scans, the need to initialize the ICP process is indeed eliminated, but the computational effort and the volume of the registered data is significantly increased. Moreover there are cases e.g. a sphere, a rotational surface such as that of a cooling tower or a pot, where the object's surface is such that, if there is high correlation of the overlapping point-clouds, the ICP falls at a local minimum and the point of convergence does not correspond to a correct solution. On the other hand, if sparse scans are acquired, especially for free-from objects, such as sculptures, there are cases where the selection of conjugate points is a really elaborate process and quite often it is impossible to correctly identify three pairs that will enable the initialization. In such cases scans cannot be included into processing and this may result to an incomplete model.

With the aim to overcome such problems and automatically initialize the ICP process, i.e. calculate the approximate transformation between any two given scans of an object, various approaches are suggested in literature. (Hansen, 2006) suggests a method that segments the point clouds into planar elements and by matching planes in the object space obtains a coarse registration automatically. This method is ideal for scenes that are dominated by planar structures such as built-up areas. Another approach suggested by (Makadia et. al., 2006) involves the creation of Extended Gaussian Images (EGIs) for the scans to be registered. The solution is derived by the correlation of the EGIs in the Fourier domain using spherical and rotational harmonic transforms. This method works especially well for cases of small overlap and for various object sizes. (Bae and Lichti, 2004) propose a method based on the calculation of geometric primitives such as surface normals, curvature, change of curvature etc, that remain invariant under rigid body transformations. (Biswas et al., 2006) match and register point clouds by describing them by isosurfaces decomposed into spherical harmonics. However this approach is applied for full (not partial) models and is better suited for database retrieval applications. Finally, (Gelfand, 2005) proposes a method that is based on feature points automatically selected according the uniqueness of a descriptor value.

Although the methods described in the above suggest solutions to the problem of ICP initialization, they are either based on characteristics such as planar areas, curvature change, detection of features, feature uniqueness or require that a significant part of the whole object (if not the entire object) is included in each scan. Thus none of these methods enables the registration of partial scans of a free-form object that may contain no characteristic points. The proposed approach, based on a suitably designed data acquisition strategy, can ensure the registration of scans that contain no targets by applying a very simple algorithm so as to initialize ICP.

As mentioned in the introduction, an approach similar to the one presented in this paper has already been applied for the case of a cooling tower (Ioannidis et. al, 2006). In specific, the survey of the cooling tower included 7 scans for the outside surface, and for the inside another 20 scans (10 scan-couples) that were acquired from 4 positions at 10 different directions (i.e. 10 set-ups) as it can be seen in Figure 1. A scan-couple is a set of two scans that are acquired from the same set-up by vertically rotating the scanner head so that the base scan contains at least 3 reflective targets and the top scan has an overlap of at least 30% with the base scan. Considering that the tower is 97m high with an 83m basal diameter, the scanner used for the survey of the inside was a Cyrax HDS2500™ with a FOV of 40°x40° and that it was impossible to place targets directly on the surface of the shell, it is obvious that the acquisition of mark-less scans in this case could not be avoided. The scan point interval was ranging from 6cm to 8cm on the tower's surface and approximately 22,000,000 points were acquired in total. In the absence of targets inside for the upper part of the tower the registration of scans of the was considered to be a challenge as the tower surface is unvarying and the selection of homologous points that would normally result to a good registration was impossible.

The fundamental idea that solved the problem is the estimation of the vertical rotation between the scans of a scan couple and the application of a rotation that restores the mark-less scan to its approximate correct location. Figure 2 illustrates the results of the method described for one of the mark-less scans. Notice in yellow the annotations of the targets that were used for the registration of the scans of the inside of the tower. Further details regarding the application of the cooling tower are considered to be out of scope. However it is worth noting that the method described in (Ioannidis et. al, 2006) enabled the initialization of the ICP and resulted to an average alignment error of 5mm.

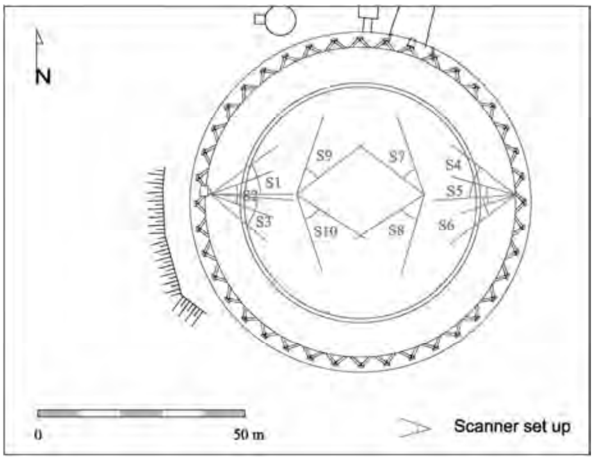


Figure 1: Plan view of the scan set-ups used for the acquisition of the inside of the Cooling Tower

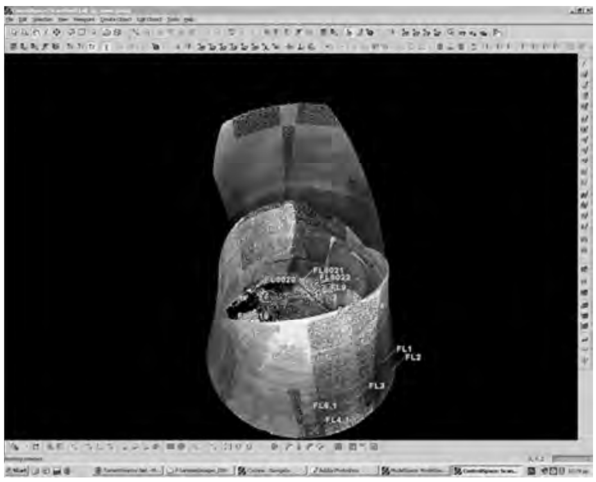


Figure 2: Example of the results of the proposed method for one of the mark-less scans acquired for the upper part of the inside of the Cooling Tower

3. PROPOSED METHODOLOGY

The main idea is to develop a simple and general method so as to efficiently enable and facilitate the integration of markless scans into the registration process. This is achieved by constraining the data collection process, i.e. scans are always acquired in couples including a first scan that contains at least three targets, so as to enable registration to a local reference system, and a second markless scan acquired by rotating either vertically or horizontally the scanner head. In this way, if the single vertical rotation between the two scans is defined, the markless scan may be rotated so that the two scans are brought in approximate alignment and the ICP can be initialized. In order to automatically obtain an approximate value of the unknown vertical angle, the space of the unknown parameter is sampled and a function that estimates the point cloud distance is calculated for every value provided by the sampling sequence until a sign change of the calculated measure is detected. A change of sign means that the interval that contains the solution that minimizes the distance measure is detected. Based on the results of the sampling process the approximate value is derived. It must be noted that in order to obtain a good approximation an overlap of at least 30% is required. In the following, practical considerations concerning data acquisition are described along with the theoretical aspects of the proposed algorithm and the technical details for the implementation.

3.1 Sampling strategy

Provided that a single rotation is the most significant part of the unknown transformation between two scans to be registered by the ICP algorithm, the initialization of the process is a rather simple problem. Let us consider that the unknown rotation is a vertical rotation, i.e. the unknown parameter is the ω rotation around the X axis. Moreover, if the equipment used is a scanner such as the Cyrax HDS2500™ that has a 40^ox40^o FOV, the unknown angle cannot be smaller than 0^g or greater than 50^g. The closed interval of $[0, 50]^g$ of the space of the unknown parameter is initially sampled with a coarse step of 5^g. For each value ω_i that the angle ω takes according to the sampling sequence, a measure of the distance md_i between the two point clouds is calculated. If there is a sign change of the distance measure calculated, according to Bolzano's theorem, this means that the unknown rotation lies in the interval $[\omega_{i-1}, \omega_i]$. This interval is re-sampled with a fine step of 1^g and the process already described is repeated. Once more, when a sign change is detected for two values ω_{i-1} and ω_i, the respective distance measures md_{i-1} and md_i are used so as to obtain an approximate value for the unknown rotation using linear interpolation (Eq. 1).

$$\omega_o = \omega_{i-1} - \frac{md_{i-1}}{md_i - md_{i-1}} \tag{1}$$

where: ω_o = approximate value of the unknown rotation
ω_{i-1} = previous sample value for unknown rotation
md_{i-1} = value of distance measure for ω_{i-1}
md_i = value of distance measure for ω_i

3.2 Point-cloud distance estimation

In order to compare two point clouds and estimate how close they are for a given vertical rotation ω_i, the following processing is done:

- The markless scan is rotated by ω_i, which is given by the sampling process, according to Eq. 2

$$Au2 = R\omega_o\, Au \tag{2}$$

where: $R\omega_o$ = rotation matrix calculated for ω_o
Au = the original markless scan
Au2 = the markless scan rotated by ω_o

- The overlapping area between Ad (the scan that contains the targets) and Au2 (the rotated markless scan) is calculated and a grid on the XY reference plane is created with a bin size that equals 5 times the scan resolution. This way a 4% sample of the points of the overlapping area is used for the estimation process.
- Ad and Au2 are processed so as to select only the points within the overlapping area. Following, a 2D delaunay tessellation is created for the selected points of each point-cloud.
- Using the 2D tessellations a search is performed so as to find the points of Ad and Au2 that are closest to the nodes of the XY grid. Regarding the points found, only those that are within a radius equal to the scan resolution are kept.
- Using the points derived from the previous processing stage two matrices are created, Zd and Zu for Ad and Au2 respectively. These matrices contain the Z values of the points of Ad and Au2 that were found to be closest to the nodes of the XY grid.
- By subtracting Zd from Zu a new matrix DZ is calculated. Through this matrix an estimation of the distance between the two point clouds, *mz,* is obtained as the median of the values of DZ. It must be noted that, ideally, one would select the mode of the calculated distances DZ. However, in order to achieve that, further process would be required so as to effectively bin the DZ values and obtain a close approximation of the mode. Furthermore, the mean value cannot approximate the mode because the distribution of the calculated distances DZ may be skewed. This happens because point-clouds usually contain non-Gaussian noise (e.g. points that do not belong to the object's surface). Therefore, the median is selected as it can be computed in one step and it can effectively represent the majority of the points that are used for the estimation of the distance of the point-clouds.

3.3 Implementation details

The algorithm described in the above is relatively simple, its execution time is only a few seconds and its main advantage is that the user does not have to define many parameters. The only parameter needed for the algorithm to run, is the scanning resolution in meters.

The approximation obtained is considered to be satisfactory as the respective transform approximately aligns the two scans and enables the initialization of ICP. An attempt to obtain a better approximation of the unknown parameter results into no improvement because the actual transform also includes small φ and κ rotations and small translation components that are calculated by the ICP process. This happens mainly because there is a difference between the physical centre around which the scanner head is rotated and the actual origin of the system of the scanner.

The algorithm as described operates for vertically rotated scans. If there is a horizontal rotation between two scans the only part that needs to change is the rotation matrix by which the markless scan is rotated. All else remains the same. Moreover tests with actual data have shown that the algorithm can be applied not only for scan-couples but also for scan-chains where many scans are sequentially acquired by rotating the scanner head either vertically or horizontally each time.

In the following section experiments conducted so as to assess the performance of the proposed algorithm are described and the results derived are presented. In section 5 the results of the application for the monument of Zalongon are presented.

4. METHOD VALIDATION

In order to validate the results produced by the proposed method an extended experiment has been designed. The data set used for the experiments was acquired during the survey of the monument of Zalongon and it includes 2 scans acquired from one set-up and 5 targets common to both scans. The two scans were acquired from the same position and only a vertical rotation of the scanner head was introduced so as to obtain the second scan. For this section the first scan will be considered as the reference scan whereas the second will treated as a mark-less scan that needs to be registered. These two scans have been processed both in Cyclone and in Matlab. This is done so as to benchmark the algorithms developed in Matlab for every part of the process against those applied by Cyclone, and ultimately create a standalone set of functions that produce reliable results. The process of the evaluation can be analyzed in two stages: 1) Data-set analysis and zero-values and 2) method evaluation.

4.1 Data-set analysis and zero-values

The first step, before any kind of processing takes place, is to calculate the distances of the targets between the original scans so as to have an estimate of the initial point-cloud configuration. The results of Table 3 reflect the fact that the two scans have been acquired from the same position with a vertical rotation of the scanner head (an unknown ω angle).

Also, before applying the proposed processing approach, the parameters of the unknown transform were calculated both in Cyclone and Matlab using all of the targets. The remaining error vector was also calculated for each target and for both cases. This step is essential so as to obtain a good value for the parameter that is defined by the proposed algorithm (i.e. the unknown ω angle). It is also useful to know how well the targets are registered once complete transformation is recovered.

As it can be seen from Table 4, the results indicate that both Cyclone and Matlab produce almost identical results with differences of the order of 0.1 mm for the remaining errors. Moreover it is derived that the unknown angle ω, in this case, is approximately 10.85 grad. As for the other parameters of the transformation it is found that the parameters of the translation vector are rather small and that the φ and κ angles are also non-zero with φ significantly larger that the k angle.

4.2 Method evaluation

The first step here is the application of the proposed algorithm so as to find the unknown rotation (ω in this case). The results are summarized in Table 5. It is worth to note, that the approximate value derived is =10.7920^g, which is close enough to the actual angle of 10.85^g.

No transform calculated				
T (X,Y,Z)	0.0000	0.0000	0.0000	(m)
R (ω,φ,κ)	0.0000	0.0000	0.0000	(grad)
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
4	-0.0202	1.9803	-0.5524	2.0560
7	-0.0251	2.4629	-0.4710	2.5077
3	-0.0254	2.5420	-0.4685	2.5849
1	-0.0322	3.3798	-0.8934	3.4960
13	-0.0321	3.1860	-0.8455	3.2964

Table 3: Target distances as calculated for the original scans

Target registration				
Cyclone				
T (X,Y,Z):	0.0003	-0.0032	0.0106	(m)
R (ω,φ,κ):	10.8493	0.1156	0.0261	(grad)
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
4	0.0008	0.0003	-0.0017	0.0019
7	0.0003	-0.0014	0.0001	0.0014
3	-0.0003	0.0008	0.0010	0.0013
1	-0.0013	-0.0006	0.0006	0.0016
13	0.0005	0.0009	0.0000	0.0010
Matlab				
T (X,Y,Z):	0.0003	-0.0029	0.0106	(m)
R (ω,φ,κ):	10.8475	0.1173	0.0172	(grad)
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
4	0.0007	0.0003	-0.0016	0.0018
7	0.0003	-0.0014	0.0001	0.0014
3	-0.0003	0.0009	0.0010	0.0013
1	-0.0013	-0.0007	0.0007	0.0016
13	0.0005	0.0009	-0.0001	0.0010

Table 4: Transformation parameters and residuals calculated by Cyclone and Matlab using the coordinates of the targets for the registration of the two scans

Coarse sampling			**Fine sampling**		
	sign(mzi)	|mzi|		sign(mzi)	|mzi|
0	(-)	1.2349m	10	(-)	0.084m
5	(-)	0.6235m	11	(+)	0.0221m
10	(-)	0.084m	**Approximate value**		
15	(+)	0.4264m	=10.7920^g		

Table 5: Results of the process for the calculation of the approximate value of the unknown rotation

Approximate alignment				
T (X,Y,Z):	0.0000	0.0000	0.0000	(m)
R (ω,φ,κ):	10.7920	0.0000	0.0000	(grad)
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
4	-0.0202	0.0070	0.0109	0.0240
7	-0.0251	0.0113	0.0090	0.0289
3	-0.0254	0.0106	0.0048	0.0279
1	-0.0322	0.0198	-0.0067	0.0384
13	-0.0321	0.0159	-0.0016	0.0359

Table 6: Target distances after rotation of the mark-less scan by ω_o=10.7920^g

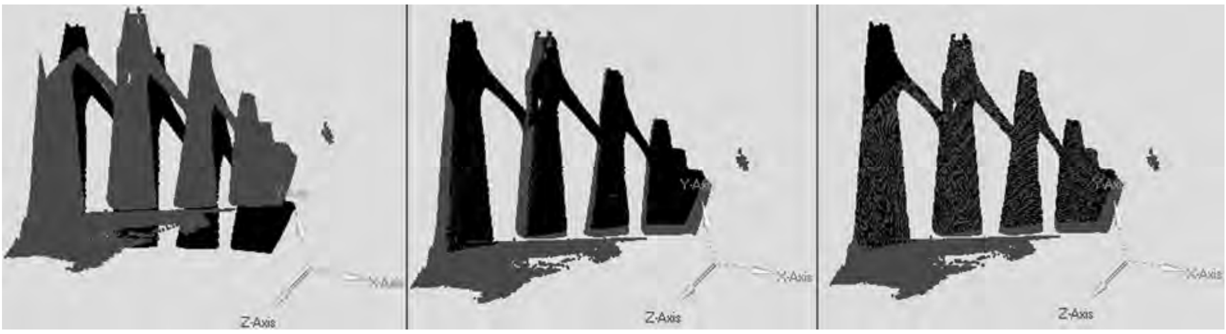


Figure 7: On the left, the reference scan (in grey) and the scan to be registered (in black) in their original positions. In the middle, the scan to be registered has been rotated by ω_o and is in approximate alignment with the reference scan. On the right, the results of the registration after the implementation of the ICP are presented.

ICP registration				
Cyclone				
T(X,Y,Z):	-0.0013	-0.0067	0.0118	(m)
R(ω,φ,κ):	10.8683	0.1080	0.0180	(grad)
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
4	0.0001	0.0004	-0.0012	0.0013
7	0.0000	-0.0005	0.0009	0.0010
3	-0.0005	0.0015	0.0020	0.0025
1	-0.0013	0.0008	0.0016	0.0022
13	0.0005	0.0023	0.0007	0.0025
Matlab				
T(X,Y,Z):	-0.0007	-0.0074	0.0122	(m)
R(ω,φ,κ):	10.8685	0.1134	0.0156	(grad)
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
4	0.0004	-0.0003	-0.0008	0.0010
7	0.0001	-0.0010	0.0012	0.0016
3	-0.0004	0.0013	0.0022	0.0026
1	-0.0013	0.0012	0.0015	0.0023
13	0.0005	0.0025	0.0007	0.0027

Table 8: Transformation parameters and residuals calculated by Cyclone and Matlab by applying the ICP algorithm

Figure 9: View of the front side of the monument of Zalongon.

After rotating the mark-less scan by the approximate ω_o rotation once again, the distances of the targets are derived. As it can be seen from Table 6, after the approximate rotation is applied, the targets are in approximate alignment indicating that the application of ICP is enabled. The distances of the targets that originally were estimated to a few meters are now only a few centimeters. The same is also suggested in Table 5 were the point cloud distance is estimated by the *mz* value (i.e. the median of the distances calculated for points close to the nodes of the grid defined within the area of overlap).

The final step of the evaluation process is the application of the ICP algorithm in Cyclone and in Matlab. The ICP is performed in Matlab by a modified version of the algorithm "*icp2*" of Ajmal Saeed Mian, of the Department of Computer Science of The University of Western Australia. The modification applied is that the $R\omega_o$ matrix is also passed as an argument in the list of arguments of the original algorithm so that every time the ICP is repeated the initial rotation is taken into account and the complete transformation is calculated. Once more, as shown in Table 8, the final results produced by the algorithms in Cyclone and in Matlab are approximately the same.

As illustrated at the left image of Figure 7, the two scans are originally in such positions that they cannot allow the initialization of ICP. The middle image of Figure 7 shows the approximate alignment after the vertical relative rotation is restored. Finally, at the right image of Figure 7 the creation of Moiré patterns ensure the high quality of alignment of the two scans by the proposed process as applied in Matlab. Figure 7 indicates that the selection of conjugate points was virtually impossible in this case. Nevertheless the results of the proposed approach for the approximate alignment of the two scans are rather satisfactory and the final results of the application of ICP are just as good as those of Cyclone. The major difference here is that in Cyclone the ICP process could not be initialized with the point clouds in their original position, but target-registration had to be performed first.

5. APPLICATION

The monument of Zalongon (Figure 9) is a 12.5m (height) by 17.5m (wide) complex of sculptures situated in NW Greece, that was built in the 50's according to the designs of Patroklos Karandinos and George Zongolopoulos. It is attributed to the women of Zalongon who according to the legend holding hands with their children danced and fell from the cliff so as to avoid getting caught by the enemy. The monument is built almost at the edge of the cliff that is 80 m higher than the nearest area that is accessible by car. In Figure 9 there is a view of the front side of the monument.

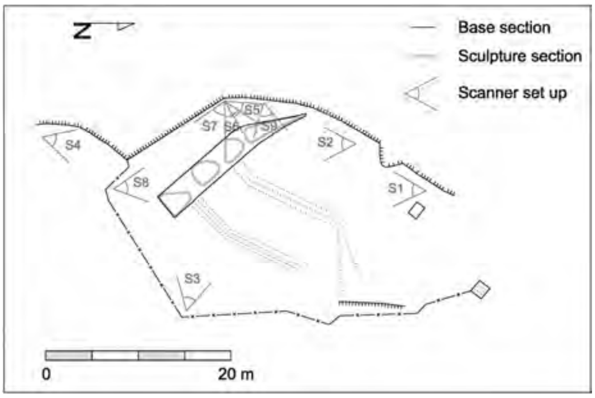


Figure 10: Horizontal plan of the monument and its surroundings including the positions of the scanner set-ups.

For the survey of the monument a reflectorless total station, the Cyrax HDS2500™ and a 10 MP camera (Canon EOS5™) were used. In Figure 10 a diagram of the scanner set-ups is presented. All of the scans were acquired from 9 scanner set-ups. However, the data acquisition for the front side of the monument was not as easy as it was for the back side. Specifically, only three single set-ups were enough to cover the back side (S1, S2 and S3), but it took 6 set-ups and a total 10 scans to cover the front side; three single scans(S4, S8 and S9), 2 scan-couple (S5 and S6) and one scan triplet (S7). The most significant problem of the survey was the limited space that was available in front of the monument as the fence was 2 - 5m away from the front side. The only case where a scan of the front side could be acquired from a greater distance was that of S4, where all of the equipment was set-up on a rock over the fence. For the rest of the scans that were acquired, all set-ups were inside the fence that was just a few steps away from the cliff. It is clear that a distance smaller than 5m is not enough to allow for a full height scan of the monument that has a maximum height of 12.5m. Therefore scans that were acquired from set-ups 5-9 were acquired sideways. Moreover, the fact that the monument is inaccessible by car and that any equipment used must be carried on the top of the cliff makes impossible the application of popular -but also expensive- solutions such as use of scaffoldings or cranes for the placement of targets. Thus targets could not be placed higher than 3m.

Taking all of these under consideration the acquisition of mark-less scans could not be avoided. However, aspects concerning the optimal and complete acquisition of the objects had to be considered. For example:

- the south-western of the sculpture forms, that is the highest of all, had to be acquired with three scans so as to guarantee that an overlap of 30% would be retained
- the north-eastern form, that is closest to the fence, was not optimally acquired by the other set-ups, so a new set-up just for that part was necessary (S9)
- scan-couples were acquired from set-ups 5 and 6.

A total of 18 reflective targets were used so as to enable the global registration of the acquired scans. Four targets were used for the back side, 3 targets were placed on the sides of the forms and another 11 targets were used for the registration of the scans of the front side. It is obvious that the abstract forms of the monument, its location and its surrounding area render it as a rather difficult and complex object to survey.

Regarding the registration process, first all of the mark-less scans were registered with the corresponding scans that contained targets by applying the proposed method and the ICP algorithm; all processing was done in Matlab. Table 11 summarizes the results for the 2 scan-couples of set-ups 5 and 6, and the scan-triplet of set-up 7. Especially for the scans of set-up 7, the proposed approach was applied first for top and the middle scan and, when those were registered, the method was once more applied so as to register those point-clouds with the base scan. Table 11 indicates that the application of the proposed method for the initialization of ICP can produce very satisfying results. This is evident from the values of the residuals that were calculated for targets visible in the various scans. It must be noted that these targets could not be used for the registration of the scans as a maximum of two targets were visible in each case but they served as check points so as to evaluate the process. For each registration the average alignment error is also calculated as the mean of the absolute distances along the X, Y and Z axes of all of the point correspondences found by ICP.

Scan couple of set-up 5 (top to base)				
ICP initialization for =25.0327^g				
T(X,Y,Z): R(ω,φ,κ):	0.0060 24.0336	-0.0092 0.4324	0.0240 -0.0561	(m) (grad)
Average alignment error = 0.0054m				
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
11	-0.0003	-0.0020	0.0016	0.0026
10	-0.0003	-0.0038	0.0012	0.0040
Scan couple of set-up 6 (top to base)				
ICP initialization for =35.6083^g				
T(X,Y,Z): R(ω,φ,κ):	0.0084 34.6002	-0.0177 1.0728	0.0313 -0.7091	(m) (grad)
Average alignment error = 0.0055m				
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
16	-0.0016	-0.0071	0.0013	0.0074
Scan triplet of set-up 7 (top to middle)				
ICP initialization for =18.1991^g				
T(X,Y,Z): R(ω,φ,κ):	0.0013 18.2504	-0.0145 0.2467	0.0224 -0.0677	(m) (grad)
Average alignment error = 0.0070m				
Scan triplet of set-up 7 (top & middle to base)				
ICP initialization for =18.1828^g				
T(X,Y,Z): R(ω,φ,κ):	0.0016 18.1147	-0.0094 0.1922	0.0190 -0.0350	(m) (grad)
Average alignment error = 0.0067m				
TargetID	X-error (m)	Y-error (m)	Z-error (m)	Total error (m)
15	0.0005	-0.0027	0.0009	0.0029
6	0.0022	-0.0009	0.0009	0.0026

Table 11: Results of the proposed method combined with the respective results of the ICP applied for the mark-less scans of the monument of Zalongon. For each registration the approximate value of the angle, the transformation parameters that ICP yields and the residuals calculated for any targets available in each case are given.

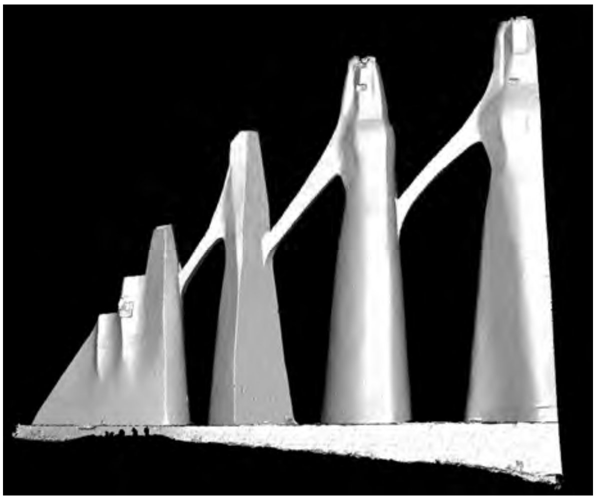


Figure 12: The unprocessed 3D polygonal model of the monument (front view).

The average alignment error ranges from 5 – 7 mm, a value rather satisfying if the technical difficulties of the acquisition process are considered (free-form object, scans acquired sideways) and the 5mm accuracy of the scanner in single point positioning is taken into account.

After all of the mark-less scans were registered with the corresponding reference scans in Matlab, the global registration of all of the point-clouds was performed in Cyclone using both target- and cloud-constraints. It must be noted that due to the object's shape there was only little overlap between the scans of the front and the back of the monument. Thus the 3 targets that were placed on the sides were especially useful because they worked as additional constraints between the corresponding point-clouds. Finally, the errors of the constraints used varied between 1 - 6mm and the mean absolute error of the overall registration was estimated to be 2mm.

All of the registered point clouds were imported in Geomagic Studio™ software and the polygonal model of the monument was created. A view of the unprocessed 3D model can be seen in Figure 12.

In order to evaluate the accuracy of the process that lead to the registration of the point-clouds it was preferred to create a polygonal model with a point spacing equal to the scan resolution and only a minimum of noise reduction was applied. The evaluation in Geomagic was performed through the comparison of the polygonal model with the geodetically acquired coordinates of 140 control points. Figure 13 illustrates the results of the comparison. As indicated by the histogram on the spectrum scale the majority of the points are within 6mm from the polygonal mesh surface. This is also confirmed by the statistics of the comparison process, as the derived standard deviation of the distances is 6mm, a value equal to the single point accuracy of the acquisition process. Moreover, the classification of the points according to the calculated deviation indicates that the accuracy of 6mm is uniform for the entire model surface.

6. CONCLUSIONS

Terrestrial laser scanners are increasingly used for applications that involve the detailed capture and modeling of various objects of interest. Especially in cases of cultural heritage and industrial applications, the acquisition process is often subdued to space limitations. For example, the object-instrument distance may be small compared to the height of the object and in many cases it may be difficult or even impossible to place targets above a given height. Consequently, data acquisition most often involves a few scans that contain no targets but need to be integrated into processing so that the final model be complete. Rather usually, these markless scans are acquired by a vertical rotation of the scanner head. However, in the absence of targets, it is often difficult to initialize the ICP algorithm. Especially for cases of free-form objects, whose surface is smooth or even uniform and no characteristic points of interest can be identified the initialization of the ICP may virtually be impossible.

The paper suggests a solution in order to bypass this problem and actually initialize the ICP process without the use of targets or conjugate points. The method involves the acquisition of scan-couples, i.e. scans that are acquired from the same set-up and for a constant horizontal direction but by a different vertical angle. The base scan must have enough targets so as to be registered-georeferenced but the top scan may be completely markless.

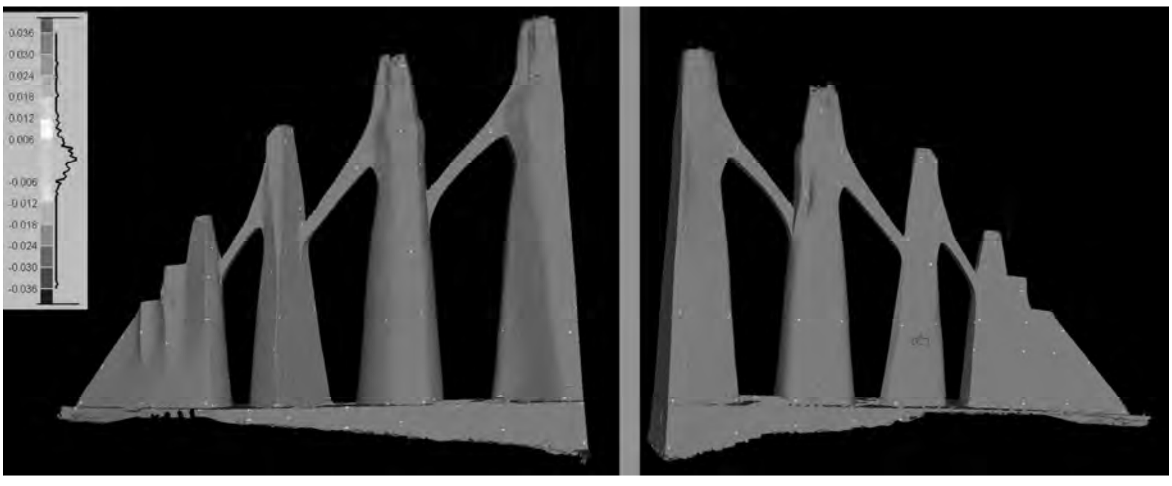


Figure 13: Results of the evaluation process for the comparison of the unprocessed surface of the polygonal model; on the left image the front view of the object is illustrated, whereas on the right the viewer sees the back side.

The proposed method has successfully been applied for the case of a cooling tower (an industrial application) and for a monument (a cultural heritage application). In both cases the proposed method enabled the integration of the markless scans into the registration process and thus led to the completion of the 3D models.

In order to validate the proposed process an extensive experiment has been implemented. The thorough investigation of the accuracy and quality of the results for each step of the process suggests that the method introduced here may be the solution for cases of free-form objects where only partial scans may be obtained and conjugate points are difficult to identify.

The proposed method as presented enables the detection of an unknown vertical angle but with a minor adjustment, i.e. the change of the rotation matrix used, a horizontal angle may be derived instead. Also, the method has been applied not only for scan-couples but also for scan-triplets. Thus another extension could involve the application for the case of scan-sequences where a few scans are acquired sequentially by rotating the scanner-head either vertically or horizontally.

The presented approach although quite simple effectively leads to the solution of a rather difficult problem, i.e. how to initialize the ICP process when there are no common targets in the area of overlap and conjugate points are difficult to recognize. The applicability of the method especially for free-form objects indicates its potential for various types of applications.

7. REFERENCES

Bae, K. H. and Lichti, D. D., 2004. Automated Registration of Unorganised Point Clouds from Terrestrial Laser Scanners. In: O. Altan (ed.), *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, Vol. XXXV, Part B5.

Besl, P. J. and McKay, N., 1992. A Method for Registration of 3-D Shapes. *PAMI* 14(2), pp. 239–256.

Biswas, S., Aggarwal, G., Chellappa, R., 2006. Invariant Geometric Representation of 3D Point Clouds for Registration and Matching. *IEEE International Conference on Image Processing*, pp. 1209 - 1212

Gelfand, N., Mitra, N. J., Guibas, L., 2005. Robust Global Registration. In: M. Desbrun, H. Pottmann (eds.), *Proceedings of the Eurographics Symposium on Geometry Processing 2005*, Vienna, Austria, pp. 197-206.

Ioannidis, C., Valani, A., Georgopoulos, A., Tsiligiris, E., 2006. 3D model generation for deformation analysis using laser scanning data of a cooling tower. In: *Proceedings of the 3rd IAG and 12th FIG Symposium*, Baden, Austria, http://www.fig.net/commission6/baden_2006/ (accessed 23 Jun. 2006).

Makadia, A., Patterson, A. I., Daniilidis, K., 2006. Fully Automatic Registration of 3D Point Clouds, *CVPR06* (I: 1297-1304).

Mian, A.S. Implementation of ICP algorithm, www.csse.uwa.edu.au/~ajmal/code/icp2.m (accessed 23 Jun. 2006).

von Hansen, 2006. Robust Automatic Marker-Free Registration of Terrestrial Scan Data. In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Bonn, Germany, Vol. XXXVI, Part 3, pp. 105-110.

8. ACKNOWLEDGEMENTS

The authors thank Ajmal Saeed Mian [ajmal@csse.uwa.edu.au, Department of Computer Science of The University of Western Australia] for the use of his implementation of ICP.

ORTHOWARE: SOFTWARE TOOL FOR IMAGE BASED ARCHITECTURAL PHOTOGRAMMETRY

A. Martos [a], S. Navarro [b], J.L. Lerma [b], S. Rodríguez [a], J. Rodríguez [a]

[a] Metria Digital, 33428 Parque Tecnológico de Asturias, Llanera, Spain - (martos, srodriguez, jrodriguez)@metria.es
[b] Universidad Politécnica de Valencia, 46022 Cº de Vera s/n, Valencia, Spain - jllerma@cgf.upv.es, sannata@upvnet.upv.es

KEY WORDS: Orthoimage, Close Range, Photogrammetry, Architecture, Matching, Software, OpenGL

ABSTRACT:

The usual photogrammetric work-flow for producing true-orthoimages in cultural heritage depends on finding and matching homologous features among different images. This process basically relies on manually choosing a reasonable model from a set of points for a subjective proper model reconstruction of the element. This is a time-consuming, repetitive and blind recognition process requiring some spatial intuition and experience from the user. Besides, either automatic image analysis techniques (mainly image matching) or laser scanning, can be used to improve this processing. But the outputs are huge point clouds with redundant information that often is not essentially required in architectural mapping (especially when the final goal is just an orthoimage). Therefore, a combination of manual and automatic techniques seems to be the ideal tool for a production environment.
We present a novel photogrammetric software tool specially designed for production of high resolution true orthoimages of architectural buildings and sites. It depends on a simple field work consisting in a few images taken with conventional digital cameras. User interaction is simplified involving intermediate projection planes and a new raster pipeline. This reduces the need to zoom, pan, and even avoids unnecessary 3D point clouds. It is progressive and designed for non-specialized users, providing intuitive methods to visually diagnose the quality of partial results. The user interface is written in C++ using OpenGL, and all geometrical calculations are parallelized and optimized for interactive performance (10^9 points/second), using both the central processing unit (CPU) and the specialized graphics processing unit (GPU). Improvements in image matching for posterior densification and KLT filtering for initial semi-automatic orientation were also tested. The software presented herein reduces orthoimage production time from weeks to just a few hours.

1 . INTRODUCTION

1.1 Background

There are many and good academic samples successfully tested of different close-range photogrammetric recording techniques used to document architectural or archaeological sites, to produce maps, plans and sometimes orthoimages. But those usually require researchers to spend considerable amounts of time and creativity; hence very few of these techniques had scaled to commercial or serial production yet.

For unqualified users and in terms of wide production, digital photogrammetry is not yet a well established technique in cultural heritage. Theoretically it should be easy to measure from a few photographs but practical projects production still requires great amounts of computer and work time.

After years using mixed techniques to produce architectural orthoimages and plans in a commercial environment, we found that most of the limitations of photogrammetry are progressively disappearing. According to our experience, field documentation procedures are already quite optimized. It can be typically done by a small team composed of one or two persons in a single day, using lightweight equipment.

It is also possible to use just non-metric, even consumer grade digital cameras achieving very good accuracy. Field calibration allows good tolerance and there is no need to record station locations or to use stereo-rigs. For small-scale projects ground control points or topographical surveying are just optional if a reference scale is taken, using a measurement tape or a handy laser range meter.

Modern field documentation procedure is today not very different from making a detailed photographic report just following some specific guidelines.

But later processing with photogrammetric software tools still remains a bottle-neck in actual productivity. Existing software in most cases inherits the design philosophy from aerial photogrammetry with little improvements. It is focused on qualified users and it is slow and has little interactivity.

Here we show the results of the development of a software tool prototype specifically designed from scratch for heritage documentation of architectural elements. We describe the proposed work-flow and justify the computer technology used. It is intended to be easy to use for non-photogrammetrists, and as automatic as possible. It is designed to be suitable for low-budget projects, allowing easy start of quick projects with just the essential requirements and progress seamlessly towards highly accurate and detailed projects. The whole user interface design rests on a hardware-accelerated graphic pipeline specifically developed for this task.

1.2 True Orthoimages

There is no doubt that presently photogrammetry and laser scanner techniques can produce accurate and detailed 3D measurements and models. But both are just acquisition techniques; direct data often result too "raw" and has to be processed later. Additionally a great amount of cultural heritage documentation projects are small and low-budget. Additionally potential contractors and users are still not much used to handle

3D documents but they are very familiar with plans and maps that they often demand as final products.

Orthoimages are a nice compromise between plans and 3D models, which can make this transition more natural, and have some of the best advantages from both sides. Orthoimages are colourful, accurate, detailed and very objective. Are also easy to show, understand and handle, and can overlay with maps and plans. Often it even results easier to produce a high resolution orthoimage first, and then drawing the plan by copying over it.

True-orthophotographs are digital images where every point of the object is orthogonally projected with a strict pixel-level accuracy. This is different from single plane ortho-rectification or other approximate techniques, where the original photos are simply corrected and projected to an idealized perfect plane that do not exactly fits the shape of the object surface.

Generation of these *true-orthophotographs* is not very usual in architecture till now due to some drawbacks of traditional methods. Sometimes they are produced by differential rectification, projecting original photographs onto a previously known three-dimensional DSM (Dense Surface Model) which could come from photogrammetry or from other sources (Tauch and Wiedemann, 2004). As a high resolution orthoimage requires a very dense model, this implies lots of work and that is why usually laser scanner is preferred (Boccardo et al, 2002). But this is a costly solution in terms of field work, and precision of the final true-orthophoto depends on the quality and detail of that intermediate DSM.

Some other photogrammetry software produces simple two-dimensional projective transformations. But just a simple rectification, or a mosaic from a small number of them, is accurate only if the object or its parts are simple or flat and shots are taken frontally enough, so that the displacement from the real projection is small comparing to the pixel size on the final representation. When using a single plane, these errors are often evident in windows, doors or other areas far from the idealized plane (Figure 1), but as real surfaces are never completely flat, smaller errors are always present everywhere.

The next logical step to improve this method and produce still approximate but better orthoimages of simple buildings (Wiedemann et al, 2003), is to idealize or "model" the rough shape of the real object by defining a simple 3D polygonal mesh using a few 3D points on the object surface. This means defining smaller additional planes for some details, like windows or doors.
That is a extended practice for simple elements, but the problem is that even the flattest real surface is not flat at all. Therefore it may be acceptable when the object is simple and resolution requirements are low, but as it increases the method becomes impractical. When projecting more than one image onto the mesh and onto the same orthogonal plane, they should look the same. But they do not fit at all at high resolution. Smaller relief differences from the mesh become quickly evident as projection errors and the pseudo-orthoimage tend to look distorted. Also it is not unique, depending on the original image used as source.

Figure 1: Simple façade. Two plane rectified images. Differences are evident at the window but smaller errors are also present everywhere.

To deal with partial occlusions the process of composing a final orthoimage from different source photographs is unavoidable and it can be an unpleasant task if different projections do not precisely match.

Additionally it is a non-trivial task for the user to choose a good polygonal "model" that is both simple, to save working time, and that reasonably fits to the real surfaces. It requires guessing and accurately marking in advance the points that later will define a good mesh. This is a blind process requiring user experience and also takes time to process.

1.3 New graphic pipeline

Traditional photogrammetric work-flow can be summarized in choosing several features (point, edges...) from one image and looking for its homologue matches on other images. Then some calculations, like image distortion correction, rotation, projection and intersection, are used to obtain 3D coordinates for each feature. This matching work is usually performed in the space of the original photographs where perspective can make the images look quite different for the user and also for automatic algorithms.

A key concept in the design of this software tool is a novel graphic interface that allows full interactivity and corrections or refinements to be done at any time providing intuitive diagnostics from partial results. This pipeline involves very intensive calculations achieved through a software development specific for GPU (Graphic Processing Unit), using *OpenGL* (Open Graphics Library).

Great effort was put on optimization using custom algorithms to handle images and matrices directly on the Graphics Processing Unit that provides huge speed-up and the capacity to process all visualization tasks in real time. Cross-correlation, orientation and camera calibration also benefit from great performance improvements and the user interface becomes very simple and efficient.

2. METHODOLOGY

This software aims to fully cover the office work stage for orthophoto production. The typical work-flow is as follows:

2.1 Image organization

Field work flexibility is very important. Despite the fact that repeated visits to the site and extra field hours can become expensive, it is often hard to plan and organize field photography for large cultural heritage photogrammetric projects. Some shots have to be taken from improvised locations due to limited access to some locations, variations in sun angle, changing illumination from cloud covering, occlusions (people, vegetation, vehicles...), etc.

We have designed a user interface allowing quick organization of images, not just at the beginning of the office work, but also early in the field being able to run on a mobile computer.

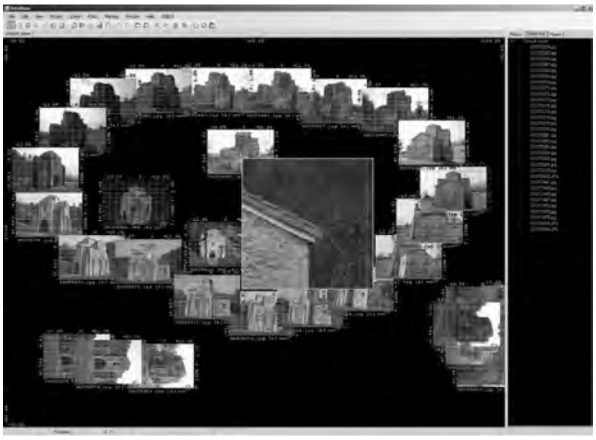


Figure 2: Desktop screen capture. Images are organized according to a spatial pattern. A magnifier permits quick inspection of details

This interface allows the user to just drag the images from disk storage or any external photo-manager software, and drop them into configurable virtual desktops. Once there, images can be interactively placed according to any criteria of choice. For instance, logical or physical organization of the project, number of common homologue marks or image quality.

To avoid memory limitations but still ensure an agile response, images that are not used for a while are automatically reduced, freeing memory for the most recently used ones. They are still promptly available in full resolution allowing quick zoom and inspection. The remaining images are loaded on demand when the user clicks on them. This was tested on an ordinary laptop with up to 400 simultaneous 16 Megapixel images with no apparent lag.

This manual organization (Figure 2) is just optional and in simple projects it can be postponed for later, if ever needed. Layouts can be multiple and as images are stored just once they can be switched at any time instantly. Later on, when external orientation parameters can be resolved, the provisional layout can naturally evolve to the 3D physical view of the model and camera locations (Figure 5).

2.2 Initial orientation

Once the project is roughly organized the next step requires the user to focus on a small set of images at a time, conveniently fitting in memory and working area, at full resolution.

Traditionally the next step would be calculating accurate relative orientation and camera positions for each image, for what a set of a few mark pairs of homologue features is needed. Afterwards bundle adjustment will calculate external matrices containing relative camera physical position and orientation for each image, allowing 3D calculations henceforth.

Early orientation is important to quickly start a new project, but the threshold of required pair marks starts from an absolute minimum of 5-8 per image, depending on the algorithm (Stewénius et al, 2006), but better 20, to handle possible mismatches and obtain a good accuracy. This requirement can consume a significant amount of time if the user has to fulfil it manually.

For proper orientation, marking has to be accurate and cover the images fully. Thus it is required either to navigate at full resolution or use a magnifier tool looking for features. This means constantly jumping between different places when there are still no helpers or constraints like epipolar lines. Often users tend to become "lazy" and avoid adding convenient or necessary new images or marks to an existing project.

This working time is reduced using a fast interface with high refresh rates. But properly choosing and marking these features is still a non-trivial task, requiring user experience and some trial and error. High contrast stickers or natural stains in flat surfaces are easy, unequivocal and accurate marks for orientation purposes, but may be of little interest to describe object geometry. Marks at edges are useful to properly model objects, but look quite different when camera convergence angle is high or if the objects are eroded. Users tend too to mark "on the air" introducing systematic bad correspondences on them, thus they should not be used for orientation.

In this software an automatic marking algorithm finds at first a few hundreds of good candidate features on each image, according to KLT feature tracker (Lucas and Kanade, 1981). Then, they are cross-correlated to find homologue matches in other images. Best relative orientation for each pair of images is determined using RANSAC algorithm (RANdom Sample Consensus, Fischler y Bolles, 1981) to discard false correspondences and determine a good set of a few homologies in short computation time (Tomasi and Kanade, 1991; Shi and Tomasi, 1994). When found these pairs are usually very accurate (1-2 pixel or better, according to our sample projects).

When a new image is first loaded, in some cases this algorithm can automatically obtain at least a dozen pairs of common features between other already oriented images. Then bundle-adjustment can establish a good initial external orientation for the new photographs with little or no user intervention at all.

If this method fails to provide the minimum of good pairs, this is easily detected in blending mode (Figure 3). Then the user might need to manually add or revise just some of them. Meanwhile, camera matrices are simply interpolated, just in case everything else fails. Right when each photograph is loaded for the first time, it is provisionally assigned with a "best guess" orientation matrix. It is obtained by interpolation from other matrices, based on the placement in the desktop working area provided by the user, i.e. if the user places the new image right between two already oriented images. This may look a too rough approximation, but the resulting matrix is often just good enough to proceed to next stages; automatic orientation or even image matching (described later).

2.3 Accuracy on orientation and calibration

Once there are enough marks, orientation is calculated "for real", but usually still roughly. When 20 marks are available, matrices are better defined, but there is not a clear limit. More than a hundred marks produce even more reliable orientation, and even better if pairs become triplets, quadruplets, and so on.

This illustrates that all relative orientation matrices are, by nature, always provisional approximations to some degree, constructed by weighting redundant (or as seen, even insufficient) marks. Consequently those matrices can be refined at any time when more homologue marks are available.

Figure 3: Portada de San Vicente. Two original images blended and overlapped for inspection. White segments join found mark pairs

Further 3D calculation such as ray intersection or image projection will greatly depend on the accuracy of these matrices, and also from camera internal model (calibration). Both orientation and calibration depend on the number of marks, their distribution over the images and object, and their accuracy. Obviously the more homologue marks, the better.

Regarding accuracy, it may look compulsory to do marking process into the original image coordinate space, as doing so on other spaces might lead to error accumulations. But as seen next, it is perfectly possible and rigorously accurate to continue marking homologies on other more convenient spaces, even without reliable orientation or calibration yet.

2.4 Graphic pipeline redesign

Our design and development depends on accelerated graphic library (*OpenGL*) to perform all point transformations in real time, including lens distortion correction and projection. For each pixel shown on the screen, many times per second and over multiple images pixels are blended at the same time. This high performance pipeline provides a few novel unique features.

Unlike in most other photogrammetry software, instead of working only into original image spaces, we can instantly switch at any time to show, navigate and mark points in other more convenient planar spaces, i.e. a rectified image space where lens distortion was removed showing straight epipolar lines, or an approximate object plane space in a fictitious surface near the object, or even the final orthophoto space.

Any update on camera orientation or internal calibration is instantly reflected in every mark feature and in every pixel shown on the screen. Although external orientation, principal point and lens distortion parameters may still be rough approximations, and therefore coordinates in new spaces will not be very accurate (Mayer, 2003), this will not pose a problem to continue working rigorously.

When new homologies are marked by the user or by any image algorithm, in any of the available working spaces, their coordinates are seamlessly traced to original image space coordinates, and stored there. Even if the projection is not exact and if the image looks distorted the marking is surely where the user intended. These new marks will provide progressive refinements of rough camera and object model, until the desired accuracy is achieved.

The pipeline also provides a very intuitive and straightforward diagnostic method for the consistence of orientation, marking and disparity. The graphic system allows accurately overlapping and blending any set of images, changing their transparency in real time using the mouse, or switching them to front or back instantly. It is then evident at a glimpse where images match precisely and where not (Figure 4).

Figure 4: Two images blended on working plane space. Doors define the plane hence looks sharper. Parts away from the plane look blurrier

2.5 Working planes

Instead of working onto original image spaces, we found out that working onto other intermediate planes or onto the final orthoimage plane makes processing easy, robust, more accurate and faster, both for the user and for automatic algorithms.

The user starts by just roughly choosing arbitrary but convenient working planes, i.e., the main façade. Original images are projected instantly onto that plane and two or more images can be blended. For clarity, this procedure is recommended for inclinations just up to 45 degrees from the working space, choosing a perpendicular working plane for the rest, but still works fine for greater inclinations. Anyway the plane can be switched at any time.

This way the user can notice in a very intuitive way which areas fit better onto that plane, and where the images do not exactly match in blending. Some features appear displaced in different positions, becoming their disparities naturally evident. This projection can be quite good for quasi-flat surfaces, but very rough approximation for others. In a further step we can define a more detailed mesh over this surface to fit it better.

The working plane also allows the minimum user effort through defining new nodes in the mesh only where seem to be needed, and provides an immediate and clear idea of how intermediate progress adjust and what has to be added or corrected next. Matching algorithms that search for homologies can also take advantage in these spaces by using smaller search windows, becoming way simpler and faster.

2.6 Marking in blending mode

Marking new homologue points pairs is a very easy and quick task to accomplish in "*blending mode*" onto a working plane where two or more images are shown overlapped. Unlike when marking pairs into separate images, where is easy to loose the track, here the user can quickly navigate and zoom at full resolution through both images synchronized.

In the chosen sample (Figure 4 and 5) the wooden doors look already sharp, meaning that they are very close to the chosen working plane. If convenient this is the best moment to quickly mark several pairs over them, with single mouse clicks for both marks in each pair (see doors in Figure 6)

If the blended image looks a bit blurry at some area, this means it is required just a small adjustment. Instead of moving the cursor, optionally one of the images can be slightly displaced until the area looks sharp again. This action is equivalent to switch to new parallel working planes and allows marking of new points at different heights. The displacement is recorded hence just releasing the mouse button creates the complete pair.

As the cursor has to travel very small distances, it is fast and intuitive to find and mark several homologies avoiding jumps between images. Using this method marking manually hundreds or thousands of points can take just a few minutes. We experienced marking rates of up to 1000 pairs of marks in less than 15 minutes.

2.7 Mesh creation

Next step to improve the adjustment is to create a polygonal mesh that approximately models the object. In this sample (Figure 6) the right part of the object shows a simple mesh with 25 marks, created in less than one minute in blending mode. This mesh is simple, but by far enough to proceed to the next stages. The left part is intentionally unadjusted for illustration purposes. In blending both images look blurry in that area making evident for the user the need for better adjustment.

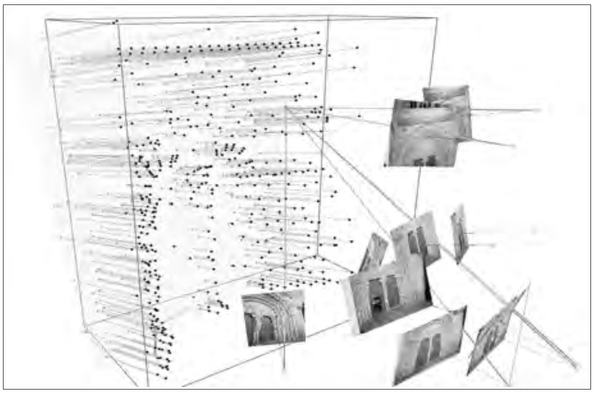


Figure 5: 3D view of the working plane used and relief errors

Instead of projecting both blended images into a well defined three-dimensional single plane, there is an alternate short-cut. Using the mesh, one image can be interactively shown "deformed" until it matches another. This has the advantage of being possible even when there is no orientation information at all and with a provisional 3D plane.

Using the deformation mesh there is not too large difference between cameras being calibrated or not, or if the features used to orientate are not accurately located at first. It is simply enough that the mesh makes the two images to visually "match" in the areas of interest. If they do, the nodes of the mesh are new pairs, accurate and reliable as they are tracked back to original coordinates.

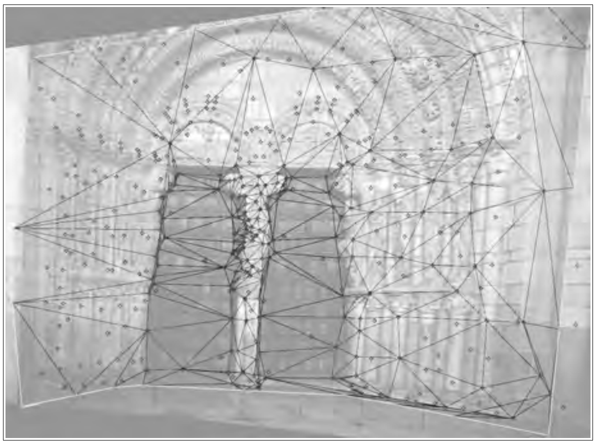


Figure 6: Mesh used. Left columns are roughly modelled by a flat mesh with no interior marks. Right side uses a mesh with just 25 points

This is a robust method, largely tolerant to errors. It provides a natural way to diagnose errors by simply watching closely areas where two or more images does not fit or look blurry. New homologies and/or new mesh nodes are created right where they are more necessary, reducing to a minimum the number of nodes.

In each step as more homologies become available they can be used to refine orientation and calibration, which in turn finally will refine projection for marking the next pairs. The user can decide to finish the process at any time depending on the immediate feedback seen on the resulting projection images in the orthoimage plane or in 3D view. When different images already match to the desired accuracy, no more work is necessary (Figure 7). Even if every mark is marked manually, the process takes the minimum working time.

Figure 7: OpenGL "near orthophoto" 3D view. Geometrical depth noise is less noticeable in this view therefore the mesh is adequate

2.8 Disparity map

For simple elements, former steps may be enough to produce good quality orthoimages and even simple but nice 3D models, but further steps are needed to reach pixel level accuracy when surfaces are complex and very detailed.

After images already roughly match in the working space, a fine detail disparity map is calculated using image correlation (Figure 8). This map provides the remaining small corrections and produces millions of new homologue points with some noise that is filtered later. Again some of these marks, the more reliable of them according to image structure, level of correlation and epipolar constraints can be used later to refine and update orientation and calibration.

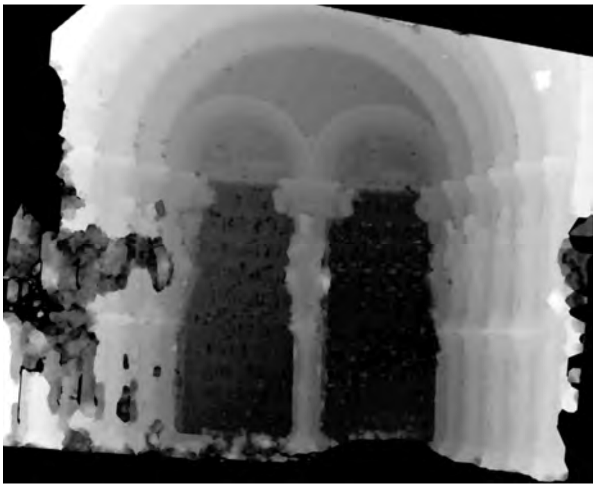


Figure 8: Fine detail disparity map (without filtering noise). Left part still requires just a few more manual marks to improve the mesh

The algorithm developed is a variation of the pyramidal correlation approach, with small search windows, optimized for parallel processing with a filling rate of 2-3 seconds per million points to provide immediate feedback and unconstrained to deal with projection errors. A sample of this map is shown unfiltered in Figure 9, where the reference working plane corresponds to the doors. Dark means low disparity, i.e. the area is more or less near the working plane, and bright means that is far from it. This information is closely related to relief as appreciated in the image hence it is easy to detect failure areas.

Image B can be corrected by this map and projected into other image space A producing a synthetic image SBA (Figure 9). Overlapping images A and SBA the matching is evident. Switching them is an immediate visual test of the quality of the map as errors are easily noticeable as residual displacements.

Figure 9: Synthetic image SBA from disparity map. Overlaps almost exactly with image A, but image information comes from image B

2.9 Orthoimage space and 3D model

After filtering and adjusting, the two original images are finally re-projected onto the real orthophoto plane (or 3D model), transparently computing intermediate corrections on working plane, mesh and fine detail disparity. Eventually, the two imagery match together providing a validated portion of the final orthophoto.

Figure 10 shows a coloured 3D point cloud, similar to those obtained from laser scanner, but just from the disparity map of the two images. It is shown unfiltered and unconstrained just for demonstration. Geometrical noise is evident in some areas and the cloud is sparse. Of course, filtering the cloud greatly reduces some of this noise and it is also possible to mesh the cloud and form a dense surface model (DSM) and even join multiple portions to increase coverage. These are well known issues in laser-scanning, out of the scope of this paper.

Traditionally each point in the images would be projected onto the DSM and then projected again onto the chosen orthogonal plane. The quality of these calculations depends critically on good camera calibration and orientation, accuracy in marking and good reconstruction of the 3D points and also a correct definition of the orthogonal plane, that in turn usually depends on everything else. Also a 3D model can be seen from any angle, therefore the mesh has also to deal with occlusions and coverage. Joining portions require seamlessly mixing the "textures" and usually reducing the model back to a reasonable amount of polygons. This is a complex process consuming memory, processing and user time and requires specialization.

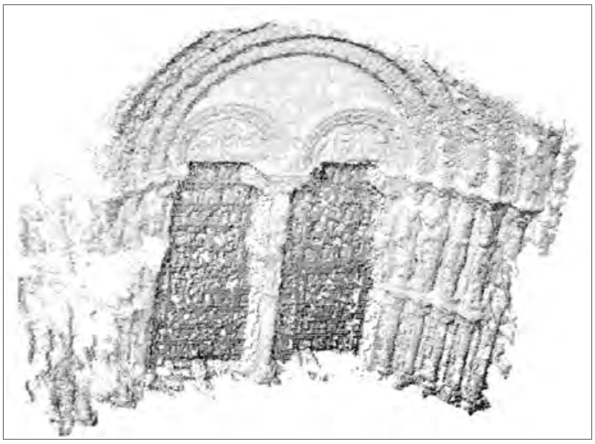


Figure 10: Unfiltered and unconstrained 3D disparity point cloud

But if the goal is just to produce true orthoimage instead 3D, production of that error prone detailed intermediate DSM is somewhat unnecessary. We can again take a short-cut and directly project the images right onto the orthogonal plane as soon as it is first defined by a few marks. This projection is not planar, it requires some 3D corrections that can be calculated right from disparities by means of a GLSL "shader".

Again switching from different source images in blending mode, if they already match means that the deformation model fits with the orthogonal plane re-projection and thus any of the projected images is already a perfect orthoimage in that area. On blurry areas because on the relief, the 2D mesh can be used again to interactively distort one image into another until fitting. This can be converted to a soft disparity map used as starting point for next correlation levels, improving matching.

2.10 Final image composition

To produce final high-quality orthophotos for printing, former operations must be repeated using different pairs of images to deal with coverage, occlusions or shading problems. Each of this partial orthoimages constitutes a different image layer to be exported to image processing software reducing to 2D image processing from there on.

Those layers partially overlap and match precisely onto the orthoimage plane, thus the task reduces to trim evidently wrong areas, revealing the next layer, and perform radiometric equalization. If there is more than one image covering the same area (and this is usual in a photogrammetric record) it does not matter metrically which one is used as source for the final orthoimage rasterization. The choice can be done according to visual image quality factors or based on additional layers such as depth, disparity or image structure. It is planned to use texture blending from viewing angle and image resolution in model space (Petsa et al 2007).

Regardless of the working spaces, each layer is rasterized oversampled onto the orthophoto plane just once, at the desired resolution, using the same raster pipeline and directly back from the original images with no intermediate rasterizations. If original images on disk are processed in any way (re-sampled, colour adjusted, etc.) a new rasterization takes just a few seconds to be reload.

3. SOFTWARE AND HARDWARE

3.1 Graphic pipeline design

Orthoimages suitable for large-format printing are digital images requiring a resolution of 150-250 pixel/inch (Figure 11). For the software to be interactive, this typically means managing 2-4 overlapped semi-transparent several mega-pixel images, that have to be real-time refreshed on the screen. Every single pixel has to be corrected from lens distortion and perspective, and displaced to match its homologue on the other photographs.

For a good design it is critical that the whole image transformation is calculated in less than a fraction of second and repeating for more than one image at the same time.

Figure 11: Printing quality final orthoimage, seamlessly composed from five different layers after radiometric equalization

Being performance so important, all photogrammetric and geometrical algorithms have been first prototyped, tested and rewritten in C++ with performance in mind, using a combination of central processing unit (CPU) and the specialized graphics processing unit (GPU).

Drawing two overlapped mega-pixel images on the screen with all the required processing steps from original image to true orthoimage projection requires, for every single pixel:

1. Lens distortion elimination (non-linear).
2. Projective transformation.
3. Mesh intersection.
4. Fine disparity corrections.
5. Projection onto final orthoimage plane.

Drawing so high resolution images is not an easy task even if they were just directly displayed, but it is even harder if this kind of intermediate 3D processing is needed. Graphics card acceleration is obtained using OpenGL and GLSL.

OpenGL (Open Graphics Library) is a standard multi-platform library for 2D and 3D graphics developed by Silicon Graphics. The library offers the ability of GPU (Graphic Processing Unit) programming with the GLSL language (OpenGL Shading Language).

For critical tasks, our choice was to use the GPU (Graphic Processing Unit) located in the modern graphic cards. They can process several pixels (16 to 180) in parallel at the same time. It is also a 4D vector processing unit, making many vector operations naturally in one "clock step".

3.2 Interactivity

In general, graphics applications work in an iterative loading-processing-loading procedure. The images need to be first loaded in system memory and moved to video memory to be shown on the screen using the graphics library with hardware acceleration. The movement is slow and duplicates memory usage. Then calculations are performed in CPU in a pixel by pixel sequence. This process is very slow because the CPU is not optimized for pixel (image) calculations. Finally, the resulting image data has to be moved again to video memory using *Windows* GDI (Graphics Device Library) so the user can see the resulting transformation.

The solution for efficient real-time graphics transformations is programmable hardware GPU. This new approach offers the ability to process all the calculations in the graphics card, accessing the video memory directly. Furthermore, the GPU is designed specifically to work with images and algebraic calculations, so many of the vector and matrix operations become native, and it also has a concurrent and parallelized design which provides the ability to perform the operations in many pixels at the same time. With this GPU approach it is possible to avoid the load-process-load stages, because only one initial load is needed.

Combined with the specific design of the GPU for working with images and algebraic calculations, our initial tests showed speed-ups of 1000x in processing time on consumer grade (100€) graphics card (*NVidia Geforce* 8500).

Another fact for using a GPU system for the project, is the rapid evolution that this kind of hardware is experiencing, provided by the video-game industry. Due to it, GPU are evolving more quickly than CPU processing.

Other more complex algorithms like disparity calculations are still being tested in CPU before being ported to GPU. To keep interactivity the application runs this kind of calculations in background avoiding to stall the program interface (the hourglass cursor that other programs show when they are processing) thus the user can continue working. When it finishes, results are shown immediately on the screen.

4. CONCLUSIONS

This paper shows some early results on high interactive true ortho-imagery reducing user intervention to a minimum. *Orthoware* is a very easy to use computer program capable of fully processing a total of several billion (10^9) points per second in a conventional computer system. This is just departing from photographs taken with ordinary digital cameras in affordable field work. Office production time may reduce to a few hours (or even minutes) instead of weeks. Accuracy is self-evident in form of blurry areas that are easy to correct just by making small adjustments until they become clearer.

Orthoware comes out after several years of software development, bearing in mind commercial production. The novel characteristics of its interface are only possible thanks to the standardization of recent affordable high performance graphics systems for video-games industry.

At the time of writing this paper, the software and the methodology are still in development, being tested in parallel with data coming from current production projects that are being executed. This development will simplify the overall production of high quality orthoimages in the field of architecture, opening new ways for non-specialized users and helping to successfully consolidate ortho-imagery in cultural heritage documentation.

5. REFERENCES

Boccardo, P., Dequal, S., Lingua, A., Rinaudo, F., 2002. True digital orthophoto for architectural and archaeological applications. GISdevelopment.net. http://www.gisdevelopment.net/Application/archaeology/general/archg0002.htm (accessed 28.Sep. 2007).

Lucas, B.D. Kanade, T., 1981. An Iterative Image Registration Technique with an Application to Stereo Vision. *Proceedings of the 7th International Joint Conference on Artificial Intelligence* (IJCAI '81), pp. 674-679.

Mayer, H., 2003. Robust Orientation, Calibration, and Disparity Estimation of Image Triplets. *Lecture Notes in Computer Science*, Vol. 0302, pp. 281-288.

Petsa E., Grammatikopoulos L., Kalisperakis I., Karras G., Pagounis V., 2007. Laser Scanning and Automatic Multi-Image Texturing of Surface Projections. *XXI International CIPA Symposium*, 01-06 October, Athens, Greece.

Shi, J., Tomasi, C., 1994. Good Features to Track. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593-600.

Stewénius, H., Engels, C., Nistér, D., 2006, Recent developments on direct relative orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(4), pp. 284-294.

Tauch, R., Wiedemann, A., 2004. Generation of Digital Surface Models for Architectural Applications with Archimedes3d**.** *IAPRS*, Vol.XXXV, Part B5, pp. 350-353.

Tomasi, C., Kanade, T., 1991. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132, April.

Wiedemann, A., Moré, J., Tauch, R., 2003. ARCHIMEDES3D- An Integrated System for the Generation of Architectural Orthoimages. *International Archives of Photogrammetry and Remote Sensing*, XXXVIII, pp. 554-558.

EVALUATION OF A SIMPLER METHOD FOR LARGE SCALE DIGITAL ORTHOPHOTO PRODUCTION

A. Georgopoulos, S. Natsis

Laboratory of Photogrammetry, School of Rural & Surveying Engineering, NTUA, Greece
drag@central.ntua.gr

KEY WORDS: Orthophotography, Point Cloud, Terrestrial Laser Scanning, Open Source Software

ABSTRACT:

Digital orthophotography has been a powerful tool for large scale geometric documentation of monuments for over a decade. It conveys quantitative, i.e. metric, as well as qualitative information, so valuable for the integrated restoration studies. However, several factors prevent the smooth implementation of commercially available software for the production of digital orthophotos. This implies that the Geometric Recording of Monuments at large scale, i.e. larger than 1:100, presents several difficulties and peculiarities, which call for special attention by the users. The extremely high number of points produced using Terrestrial Laser Scanners on the object's surface, which replaces the need for producing the Digital Object Models (DOM's) from carefully taken stereopairs, may be suitably exploited to free the user from severe constraints in positioning the camera. In this paper a novel and simple method developed for the production of orthophotography at large scales is described and assessed. This method successfully produces orthophotos at large scales using a point cloud and freely taken pictures of the object, thus achieving two goals. Firstly the user may work independently from the practical constraints imposed by the commercially available software and secondly there is no need for specialized knowledge for implementing complicated photogrammetric techniques, or specialized photogrammetric or pre-calibrated cameras, since self-calibration may take place, thus making the method attractive to non-photogrammetrists. The described algorithm has been implemented using open source software and is released under the terms of the GNU General Public License (GPL) in order to enhance its wide applicability, its low cost and its flexibility, as modifications may be possible by third parties according to future needs. The algorithm has been applied to a couple of cases of geometric monument documentation with impressively promising results. A thorough test has been performed using several check points and the results are presented and discussed.

1. INTRODUCTION

1.1 Orthophotography

The photogrammetric textured representations, using mainly orthophotomosaics, have been proved to be powerful products of documentation as they combine both geometric accuracy and visual detail. They convey quantitative, i.e. metric, as well as qualitative information, so valuable for the integrated restoration studies. However, there is still an unsolved problem with the complete orthoprojection of complex objects though, as the description of the analytical shape of the object cannot always be accurate, especially in areas where points with the same planimetric coordinates show different heights. Regular grids integrated by break-lines and DSMs (Digital Surface Models) are the most popular and investigated solutions used to build-up a mathematical shape description of such an object. In both cases complex algorithms and expensive computation times must be used before and during the orthophoto production.

Although their production has already reached a high level of maturity as far as aerial images are concerned, for the Geometric Recording of Monuments at large scale, i.e. larger than 1:100, it presents several difficulties and peculiarities, which call for special attention by the users. Consequently, orthophotography is not yet fully accepted by the user community for applications related to geometric documentation of cultural heritage monuments. Architects and archaeologists are still reluctant to concede working with orthophotographs instead of the traditional vector line drawings. As a consequence orthophotography usually is not included in the standard specifications of the geometric recording of monuments. The situation is becoming worse due to the need for special instruction for planning and executing the photographic coverage to face the problems of orthophoto production for the monuments at large scales (i.e. ≥1:100). The major of such problems are (Mavromati et al., 2002a, 2002b and 2003): (1) Large elevation differences compared to distances between the camera and the object, (2) Presence of "vertical" surfaces, i.e. surfaces parallel to the camera axis, (3) Convergence of camera axes, often due to space limitations, (4) Failure of automatic DTM (Digital Terrain Model) production, as all available commercial algorithms are tailored to aerial images, (5) Necessity for large number of stereomodels in order to minimize occluded areas, (6) Difficulty of surveying convex objects.

For the first two problems special measures should be taken during both field work and processing of the data. They are the main source of practically most difficulties encountered in producing orthophotographs and the relevant mosaics. The elevation differences call for elaborate description of the object's surface, in order to allow for the orthophotography algorithm to produce accurate and reliable products. Usually, problems due to the image central projection and the relief of the object (e.g. occlusions or complex surface) can be solved by acquiring multiple photographs from many points of view. This may be compared to the true orthophoto production for urban areas (Baletti et al., 2003). However, processing can be seriously delayed for DTM generation requiring possibly intensive manual interaction or even a complete failure to produce a reliable model.

1.2 Terrestrial Laser Scanning

The production of orthophotographs presents even more special problems, as it usually is a case of a highly demanding true orthophoto. Special techniques have been proposed in the past to address these problems in the best possible way. However no clearly defined solution to the above has been implemented. Terrestrial laser scanning techniques have helped the situation a lot, as they are able to provide a more detailed description of the object's surface, a fact which contributes to a more successful implementation of traditional orthophoto production algorithms. However, even this solution is by no means complete, as it imposes limitations to the orientation of the stereopairs, the position of the projective planes for the final orthophoto and, of course, to the completeness of the final product.

The appearance of terrestrial laser scanning has already shown promising contribution in overcoming such problems (e.g. Barber et al., 2002; Bitelli et al., 2002; Drap et al., 2003; Guidi et al., 2002) and also confronting other similar applications (Baletti & Guerra, 2002). The volume of points and high sampling frequency of laser scanning offers a great density of spatial information. For this reason there is enormous potential for use of this technology in applications where such dense data sets could provide an optimal surface description for applications of archaeological and architectural recordings.

The recently published literature has shown that in many cultural heritage applications the combination of digital photogrammetry and terrestrial laser scanning can supplement each other in creating high-quality 2D and 3D presentations. Specifically, laser scanning can produce the dense 3D point-cloud data that is required to create high resolution dense DSM for orthophoto generation and can be considered the optimal solution for a correct and complete 3D description of the shape of a complex object. However, a correct DSM cannot always guarantee the generation of accurate orthophotos or even acceptable photorealistic images. This is due to a number of problems such as the perspective deformations of an image and the relief of the object (i.e. occlusions). When stereopairs are taken with a certain inclination of the optical axis it can result to a different ground resolution over the image and the effect of tilt on the image geometry can cause distortions in the resulted orthophoto. In order to maintain the visualisation effect and have photorealistic models, it is possible to supplement the distorted areas of the orthophoto by texture mapping using both the image and laser data. The extremely high number of points produced using Terrestrial Laser Scanners on the object's surface, which replaces the need for producing the Digital Object Models (DOM's) from carefully taken stereopairs, may be suitably exploited to free the user from severe constraints in positioning the camera. Thus a procedure is sought to combine TLS point clouds and freely taken photography for the production of orthophotos at large scales.

2. METHODOLOGY

2.1 Previous Efforts

The method is based on a previously reported idea (Georgopoulos et al., 2005) of thoroughly and suitably colouring the available point cloud. In this paper the idea of producing an orthophoto by projecting a coloured -extended-point cloud was firstly reported. However in that case the point cloud was produced by stereophotogrammetric techniques, which increased tediousness and time necessary.

Another similar approach is the creation of a coloured 3D model comprised of triangular elementary surfaces. These are created through a Delaunay triangulation of a point cloud using the colour from a digital camera rigidly attached to the scanner, thus saving the necessity of performing orientations. Such attempts have already been reported (Dold and Brenner, 2006, Abmayr et al., 2004, Reulke et al., 2006) and led to manufacturing such devices, mainly the series LMS-Z by Riegl company (http://www.riegl.com/).

Colouring TIN (Triangular Irregular Network) 3D models from digital images irrespective of their orientation or source have also been attempted (Brumana et al., 2005), in which case the Direct Linear Transformation (DLT) was employed. As an extension to this, more images may be used for picking the right colour, through a series of selection procedures (Grammatikopoulos et al., 2004). This helps avoid selecting the wrong colour for occluded areas (Abdelhafiz and Niemeier, 2006).

Finally in the stage of experimental development is a revolutionary device, which combines information from a laser scanner and a digital camera through a common optical centre (Seidl et al., 2006). Such a device, when operational will be able to produce coloured point clouds without the necessity of any processing.

2.2 Description of the Developed Method

The developed algorithm includes the determination of the interior and exterior orientation of the image, the correspondence of the colour information from the image to the points of the cloud and, finally the projection of the coloured points onto the desired plane. It is obvious that no rigorous photogrammetric setup is necessary for the image acquisition phase. Contrary to the conventional procedure, where image tilts are of utmost importance for the quality of the final product, they play no significant role in this present case. The final projection plane may be defined at will, thus enabling the production of a multitude of orthophotos from the same point cloud.

Thus the presuppositions of the method are at least a point cloud, with suitable point density, and a digital image of the object. The algorithm firstly relates geometrically the image to the point cloud making use either of characteristic points recognized in both data sets, or pre-marked targets. A self calibration of the camera may also be performed at this stage if required, as a metric camera may not always be available. Then the algorithm assigns a colour value to each point, through a careful selection process, in order to avoid double projections, or projection and colouring of hidden points. After the point cloud has been coloured, a suitable projection plane and the size of the pixel of the final orthophoto are defined. Projection of the coloured points on the plane results in assigning colour values to each orthophoto pixel, through a suitable interpolation process. In the case that the initial density of the point cloud does not suffice for the production of a perfectly continuous orthophotography at the desirable scale, black pixels may appear in the resulting image. A basic hole-filling algorithm has been implemented in order to handle and compensate the colour for those pixels by interpolating through the neighboring pixels' colour values. In this way the final orthophotography is produced. It is obvious that the orientation of the projection plane may be defined at will, independently from the orientation of the camera axis, or axes in the case of more pictures available, since the orthophotography is derived from the projection of the point cloud rather than the differential transformation of the initial image as is the case in the traditional photogrammetric method. This allows for the production of multiple orthophotographs projected onto different planes without the need for a dedicated image or image pair for each projection plane (Georgopoulos & Natsis 2008).

3. DEVELOPMENT OF ALGORITHM

3.1 Open Source Software

The above described algorithm has been implemented using open source software and is released under the terms of the GNU General Public License (GPL) in order to enhance its wide applicability, its low cost and its flexibility, as modifications may be possible by third parties according to future needs. The algorithm has been applied to a couple of cases of geometric monument documentation with impressively promising results. A thorough test has been performed using several check points and the results are presented and discussed.

The algorithm has been developed mainly in the object oriented C++ programming language (http://zpr.sourceforge.net/), using at the same time ANSI C elements. The whole application was written within the (IDE) Code::Blocks open code environment. Finally the following open source libraries were also used: (a) GNU Scientific Language (GSL) for the necessary mathematical operations (http://www.gnu.org/software/gsl/) (b) OpenCV, distributed by Intel and used mainly for machine vision applications (http://opencvlibrary.sourceforge.net/) and (c) GetPot, a simple library for parametrizing the execution of the algorithm (http://getpot.sourceforge.net/).

3.2 Description of basic steps

As already mentioned, the current approach follows a somewhat different course than the conventional one. The orthophoto is produced through the orthogonal projection of a suitably coloured point cloud. Hence the space relation between the initial point cloud and the available digital image should firstly be determined, in order for the individual points to be assigned a colour value. The basic steps of the algorithm are (Figure 1):

- Determination of image and point cloud orientation
- Point cloud colouring by relating points to pixels
- Selection of projection plane
- Coloured point cloud rotation
- Point visibility classification depending on their distance from projection plane
- Hole filling on the resulting orthophoto

3.2.1 Data Entry

The necessary data for the execution of the algorithm are briefly described in the following:

The point cloud
Every point is described by a triplet of space co-ordinates (X, Y, Z) and another triplet (R, G, B) determining the colour. The description may be augmented by a point name, or an intensity value as recorded by the laser scanner. For enabling the various modules to work together the point cloud is stored in an ASCII format (Figure 2), which may result to large files. A suitably selected binary format would certainly improve the efficiency.

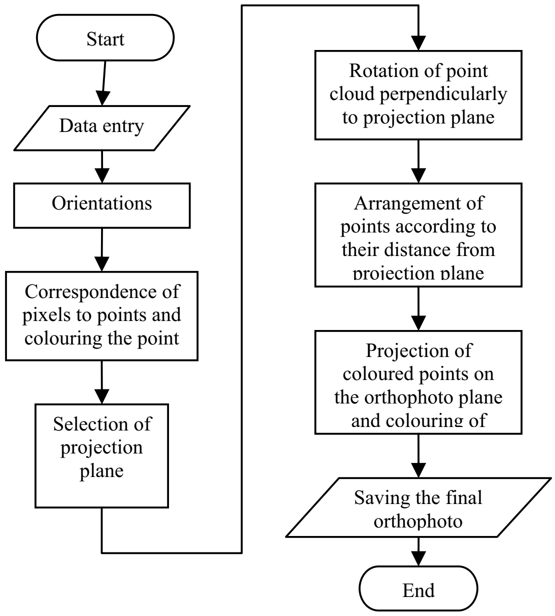
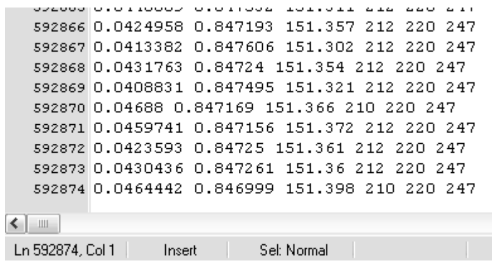


Figure 1: The flow of the basic algorithm



```
592866 0.0424958 0.847193 151.357 212 220 247
592867 0.0413382 0.847606 151.302 212 220 247
592868 0.0431763 0.84724 151.354 212 220 247
592869 0.0408831 0.847495 151.321 212 220 247
592870 0.04688 0.847169 151.366 210 220 247
592871 0.0459741 0.847156 151.372 212 220 247
592872 0.0423593 0.84725 151.361 212 220 247
592873 0.0430436 0.847261 151.36 212 220 247
592874 0.0464442 0.846999 151.398 210 220 247
Ln 592874, Col 1    Insert    Sel: Normal
```

Figure 2: Extract of an ASCII point cloud file.

The initial digital image
The necessary digital image, which will be used to assign colour values to the points of the cloud, may be stored in any of the standard digital image file formats (e.g. BMP, DIB, JPG, PNG, PBM, PGM, PPM, SR, RAS and TIFF).

Data for the image orientation
In order to determine the space relation between the image and the point cloud, which in turn will lead to the assignment of a colour value to each point, at least six control points are necessary. They are used by the orientation algorithm and may be selected either manually or digitally, with the help of two ASCII files. Moreover an initial value for the camera constant in pixels is needed in order for the algorithm to determine the interior and exterior orientation parameters, as will be described later.

Determination of the projection plane
For defining the projection plane of the final orthophoto, the user should either enter the three parameters a, b and c of the equation aX+bY+cZ=D, where D is defined by the algorithm, or alternatively may draw the straight line footprint of the plane on an orthogonal projection of the point cloud, provided that the latter is suitably georeferenced so that the Y-axis is vertical. Finally the pixel size of the resulting orthophoto should also be defined at this stage.

Point cloud management
For easy management of the points an *rgbpoint* class was declared, which contains the initial X, Y and Z co-ordinates of each point plus the additional parameters R, G, B for the colour and I for the intensity (Figure 3a). In addition, for easing the operations among the points another class *pointcloud* with basic contents a deque series was declared (Figure 3b). In this way the object oriented approach would enable the various point operations, such as rotation, translation, addition, deletion, classification and projection.

```
class rgbpoint {
    public:
        float x, y, z;
        int r ,g ,b , intensity;
};
```

(a)

```
class pointcloud {
    public:
        deque<rgbpoint> point;
};
```

(b)

Figure 3: Creation of point cloud management classes

3.2.2 Photogrammetric Orientations

For the camera calibration the algorithm developed by J. Heikkilä is used (Heikkilä, 2000). The algorithm is complemented by a model for the description of the camera-to-object relation and augmented by the radial and decentering distortion models for off-the-shelf cameras.

The mathematical model used is given by:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \propto \begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} sf & 0 & u_o & 0 \\ 0 & f & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \bullet \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \bullet \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

where $[u\ v\ 1]^T$ the image co-ordinate vector in homogeneous co-ordinates
λ scale coefficient
s the skewness factor of the pixels
f the camera constant
u_o, v_o the principal point co-ordinates
$t = [t_x\ t_y\ t_z]^T$ the translation vector from the image system to the world system
R the classical rotation matrix (ω, φ and κ)

The solution is linear and experimental results proved that an accuracy of 1/50 of a pixel may be achieved, provided qualitative image errors have been removed. The algorithm is available in MatLab (http://www.ee.oulu.fi/~jth/calibr/) and may be found in OpenCV library through the *CvCalibrateCamera2()* routine, which was used in this project.

3.2.3 Colouring the points

Collinearity condition is applied through the *cvProjectPoints2* routine of OpenCV, in order to assign colour to each point of the cloud. This routine accepts as entry the ground co-ordinates of the points and using the determined elements of the interior and exterior orientations determines the corresponding image co-ordinates. A suitable interpolation is performed in order to select the suitable colour from the digital image. The resulting R, G and B values are added to each point in the point cloud file. Points outside the image boundaries are not coloured from this image. The result of this procedure is a point cloud, which may be considered as an extended DSM (Georgopoulos et al., 2005).

3.2.4 Point Cloud Projection

The last step of the procedure is the orthophoto production, which actually is the projection of the coloured points on the desired plane. Depending on the way the plane has been defined by the user, the rotation matrix is determined for the transformation of the point cloud so that the Z axis is vertical to the plane. Not all points should be projected, as some of them may be obstructed by others. Hence the projection procedure starts from the most distant points, while the closest ones replace the previously projected points.
Based on the X and Y extremes and the selected pixel size, the size of the resulting orthophoto is determined and the transformation from X and Y values to i, j pixel co-ordinates is performed.

3.2.5 Hole Filling

Because of the projection method, it is possible to get non-coloured -black- pixels in the final orthophoto, which appear as "holes" (Figure 4).

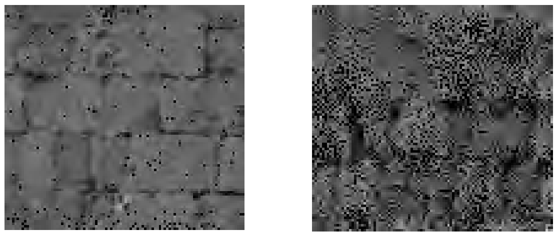


Figure 4: Examples of "holes" in the orthophoto

A hole filling algorithm is applied in order to remedy this problem. The routine *holefill()* is used for this purpose (Figure 5) and the procedure followed examines the neighborhood of the black pixel and interpolates the colour accordingly, while paying attention for cases of pixels outside the initial digital image borders.

4. IMPLEMENTATION

For assessing the performance of the developed method, a suitable object has been chosen. The Eastern façade of the Church of the Holy Apostles in Ancient Agora of Athens (Figure 6) was considered appropriate.

For the above task it was decided to produce orthophotographs both with the proposed algorithm and with well known Digital Photogrammetric Workstation (DPW). This would enable the comparison of the orthophotos and the assessment of the proposed methodology.

A Leica Cyrax 2500 time-of-flight laser scanner was used for collecting the point cloud. This scanner has a capability of a pulse laser beam of 6mm width at 50m and performs the distance measurements with an accuracy of ±4mm, with a maximum capability of 1 million points per scan. The digital images were acquired with a Canon EOS 1D Mark II, with a

CMOS 8 megapixel chip with physical dimensions $28.7x19.1mm^2$, which results to a 8.1μm pixel. This camera was equipped with a 24mm lens.

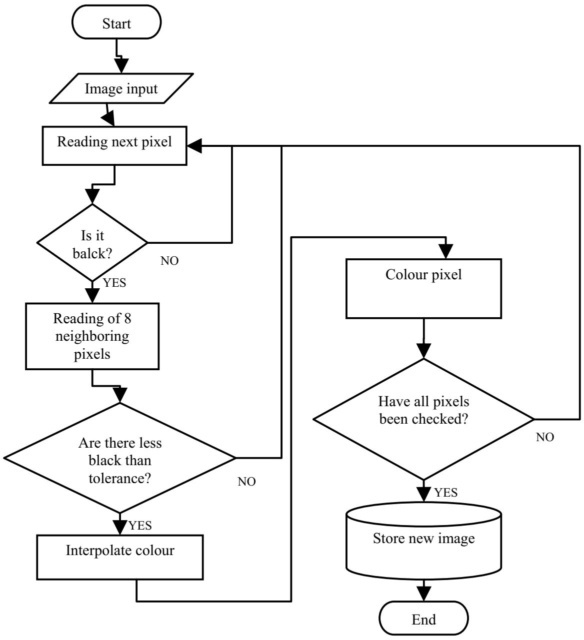


Figure 5: The flow of the *holefill()* algorithm

A point cloud was collected from a distance of approx. 10m with a density of 10mm, thus resulting to a total of 514281 points. Moreover 15 control points on the object were also measured geodetically using a Leica TCR303 total station (Figure 7). These points were premarked using the special Cyra reflective targets for accurate location and determination both by the scanner and the total station. The point cloud was processed in order to keep the points only of the area of interest, thus keeping 473736, i.e. 92% of the original points, stored in an ASCII file.

Figure 6: The Church of the Holy Apostles in Athens

4.1 Data Acquisition

The façade was photographed from various distances and various angles. "Vertical" photos were taken for use in the DPW and several additional photos were also taken in order to enable experimenting with the developed algorithm. In some highly oblique images information on the object was deliberately occluded and the results will be commented upon later.

With the above data various orthophotos were produced with different pixel sizes (5mm, 10mm, 15mm and 20mm) and using different number of Ground Control Points (GCP) (6, 7, 8, 10 and 15) for the image orientations. For the data to be operational in the DPW, some processing was necessary, as the DPW requires TIN information for the DSM. The procedure is estimated to have lasted around two hours for the production of the orthophotos.

Figure 7: Target positions measured on the façade

4.2 Orthophoto production and quality assessment

The same number and combination of orthophotos have been produced by applying the developed algorithm. In addition orthophotos were also produced using different images ("vertical" and "oblique"), but the same number of GCP's. These orthophotos could not be produced using THE DPW, as the commercial algorithm seems unable to accommodate large image tilts. Two such images, produced with exactly the same input data appear in Figure 8.

(a) (b)

Figure 8: Orthophotos produced using the same 10 GCP's and the same DSM with the DPW (a) and the proposed method (b)

The produced orthophotos with the commercial DPW are correct in the area surrounded by the GCP's. In addition their quality is not affected by the final pixel size. On the other hand the proposed method produces orthophotos correct over the whole area, even outside the GCP's. However, pixel size variation affects negatively the final quality, as expected, since the final orthoimage is directly dependent to the density of the point cloud.

Moreover, using, again, freely available image processing software (http://www.gimp.org/), differences were calculated between orthophotos produced with the DPW and the developed software using different number of GCP's. This operation reveals the consistency and strength of the algorithm employed in each case. In Figures 9 and 10 two such examples are shown. The whiter the image of the difference is, the fewer differences between the two orthophotos are present. It is evident that the proposed method consistently produces the same orthophotos both in terms of geometry and in terms of quality. Contrary, in the case of the DPW large differences are revealed, probably due (a) to the fact that the algorithms incorporated cannot cater for the peculiarities of large scale orthophotos and (b) they are strongly dependent on the position and distribution of the GCP's.

(a) (b)

Figure 9: Image differences between orthophotos produced using 7 and 15 GCP's with the DPW (a) and the proposed method (b)

(a) (b)

Figure 10: Image differences between orthophotos produced using 10 and 15 GCP's with the DPW (a) and the proposed method (b), where white denotes less differences

Highly tilted initial images produced orthophotos only with the proposed method. However the information contained in these images is distorted in occluded areas (Figure 11).

(a) (b)

Figure 11: Orthophoto produced with the proposed method (b) using initially "oblique" image (a)

4.3 Assessement of quantitative results

The metric accuracy of the orthophotos was thoroughlly examined. As in almost all cases there were some of the initially 15 available GCP's which were not used for the image orientations, their co-ordinates and distances were measured and compared. The results for orthophotos produced with the developed algorithm with pixel size 10mm are summarized in Table 1.

	#GCP's	6	8	10	15
RMS	**Δx, Δy**	10	8	3	3
(mm)	**Δs**	15	11	5	3

Table 1: Summary of quantitative results

It is obvious that as the number of used GCP's increases the resulting accuracy becomes better. At the same time the increase of GCP's beyond 10 does not appear to significantly affect the accuracy. However the requirements of a large scale (e.g. 1:50) orthophoto, even the minimum required number of GCP's (i.e. 6) returns acceptable results.

Further tests included similar measurements on the orthophotos produced on the DPW with exactly the same conditions, i.e. number of GCP's and point cloud. The comparison of the results reveals the ability of the proposed algorithm to overcome the known and, in a sense, inevitable shortcomings of almost all DPW in the case of large scale orthophoto production. IN Figure 12 the results are visualized. It becomes obvious that as the number of GCP's used for the image orientation increases, the RMS values both of co-ordinate and distance measurements decrease. However the absolute difference between the two methods in favour of the proposed algorithm is obvious.

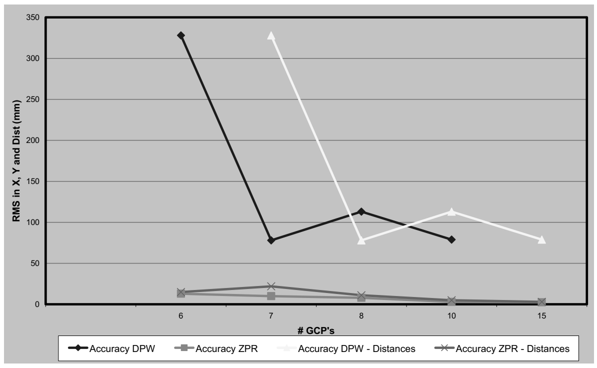


Figure 12: Performance of Co-ordinate and distance measurements on orthophotos produced by the two methods

5. CONCLUDING REMARKS

5.1 Assessment of the method

The results presented show that in most of the cases the orthophotos produced with the developed algorithm are of acceptable quality and their accuracy is within the limits of a pixel, even when the initial conditions, i.e. image orientation, DSM density, number of GCP's, are not exactly ideal, as required by commercial DPW's.

The orthophotos produced based on the initially tilted images, present similar accuracy characteristics with the others, thus freeing the data acquisition procedure from strict geometric setups. However, occluded areas still remain a problem and should be taken into account at the stage of data collection. Ideally the images should be taken approximately from the position of the terrestrial laser scanner.

The merits of the proposed method may be summarized as follows:
Simplicity and speed, as the method is fast and does not require special knowledge from the user. The procedure of orthophoto production takes about 15 minitues. In addition there are practically no restrictions in the geometry of image acquisition.
Limited need for GCP's, as the method works satisfactorily even with the least number of control points (i.e. 6).
Non-metric camera usage, as no prior knowledge of the interior orientation parameters of the camera is required, since they are

computed for every image by the algorithm with the self calibration routine.

Flexibility in producing end products, as with this method orthophotos at practically any plane may be produced, with no extra effort.

Availability and expandability, as the algorithm has been developed using Open Source software and is available for everyone willing to adapt it to his own needs.

On the other hand there are some drawbacks in the method:

- The orthophoto scale is limited by the accuracy of the scanner and the density of the collected points of the point cloud
- Terrestrial laser scanners are costly, bulky and (yet) not easily available instruments and
- The algorithm does not offer an occlusion detection routine at the stage of point cloud colouring, which results to wrong colouring in cases of steep anaglyph on the object.

5.2 Future Outlook

The most important prospects for future additions to the algorithm could be the following:

- Automatic location of GCP's, which would greatly contribute to the full automation of the process. The same result could be achieved with the automatic orientation of the image using characteristic features extracted from the point cloud, such as lines or edges, planes etc.
- Visibility control between the TLS and the image, in order to ensure the correct colouring of the points in cases of steep anaglyph, high image tilts and unfavorable taking distances. Techniques to confront this problem have already been reported (Grammatikopoulos et al., 2004).
- Using multiple images to colour the point cloud would definitely improve the resulting orthophotos, as this would practically solve the problem of non-coloured, i.e. black, points.
- The orientation of the projection plane could be such that the plane would create a section of the object, thus resulting to a multitude of more useful orthophotos for cases of occluded details on the object.
- Of course increase of the execution speed and addition of a Graphics user Interface could also be included in future improvements possible.

6. REFERENCES

Abdelhafiz, A., Niemeier, W., 2006. Developed technique for automatic point cloud texturing using multi images applied to a complex site, *IAPRS* Volume XXXVI, Part 5, Dresden 25-27 September.

Abmayr T., Härtl F., Mettenleiter M., Heinz I., Hildebrand A., Neumann B., Fröhlich C., 2004. Realistic 3D reconstruction – combining laserscan data with RGB color information, *IAPRS* XXth Congress, Commission V, Vol. XXXV, July 2004, Istanbul, Turkey.

Balletti C., Guerra, F., 2002. Laser Applications for 3D Survey of Cultural Heritage. *IAPRS*, Vol. XXXIV, Part 5, pp. 302-308.

Balletti C., Guerra, F., Lingua, A., Rinaudo, F., 2003. True Digital Orthophoto of The San Marco Basilica In Venice. *IAPRS*, Vol. XXXIV, Part 5/W12, pp. 43-48.

Barber D., Mills J., Bryan P., 2002. Experiences of Laser Scanning for Close Range Structural Recording. *Proceedings of the CIPA WG 6 International Workshop On Scanning For Cultural Heritage Recording*, September 1-2, Corfu, Greece, pp. 121-126.

Bitelli G., Capra A., Zanutta A., 2002. Digital Photogrammetry and Laser Scanning in Surveying the "Nymphaea" in Pompeii. *Proceedings of the CIPA WG 6 International Workshop on Scanning for Cultural Heritage Recording*, September 1-2, Corfu, Greece, pp. 115-120.

Brumana, R. Fregonese, L. Fassi, F. De Pascalis, F., 2005. 3D laser scanner points clouds and 2D multi-spectral images: A data matching software for cultural heritage conservation, *Proceedings of CIPA 2005 XX International Symposium*, 26 September – 01 October, 2005, Torino, Italy.

Dold, Ch., Brenner, K., 2006. Registration of terrestrial laser scanning data using planar patches and image data, *Proceedings of ISPRS Commission V Symposium 'Image Engineering and Vision Metrology'*, 25.-27. September 2006, Dresden, Germany.

Drap P., Sgrenzaroli M., Canciani M., Cannata G., Seinturier J., 2003. LaserScanning and Close Range Photogrammetry: Towards a Single Measuring Tool Dedicated to Architecture and Archaeology. *XIX International CIPA Symposium*, Antalya, Turkey, pp. 629-634.

Georgopoulos, A., Makris, G.N., Dermentzopoulos, A., 2005. An alternative method for large scale orthophoto production, *Proceedings of CIPA 2005 XX International Symposium*, 26 September – 01 October, 2005, Torino, Italy.

Georgopoulos, A., Natsis, S., 2008. A simpler method for large scale orthophoto production. *IAPRS*, 2008, 37(5).

Guidi G., Tucci G., Beraldin J-A., Ciofi S., Damato V., Ostuni D., Costantino F., El Hakim S. F., 2002. Multiscale Archaeological Survey Based on the Integration of 3D Scanning and Photogrammetry. *Proceedings of the CIPA WG 6 International Workshop On Scanning For Cultural Heritage Recording*, September 1-2, Corfu, Greece, pp. 13-18.

Heikkilä, J., 2000. Geometric Camera Calibration Using Circular Control Points, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1066-1077, Oct 2000.

Kakli A., 2004. *Contribution of Laser Scanners to the production of Orthophotographs of Large Scale*. Diploma Thesis, National Technical University of Athens, Greece, 138 pages (in Greek).

Mavromati D., Pappa P., Ioannidis C., Georgopoulos A., 2002a. Large Scale orthophotographs: Problems and Solutions. *6th Z/I User Group Meeting – Photogrammetry & Surveying* (in Greek).

Mavromati D., Petsa E., Karras G. E., 2002b. Theoretical and practical aspects of archaeological orthoimaging. *IAPRS*, Vol. XXXIV, Part 5, pp.413-418.

Mavromati D., Petsa E., Karras G. E., 2003. Experiences in Photogrammetric Archaeological Recording. *XIX International CIPA Symposium*, Antalya, Turkey, pp. 666-669.

Reulke, R., Wehr, A., Griesbach, D., 2006. High resolution mapping using CCD-line camera and laser scanner with integrated position and orientation system, *Proceedings of*

ISPRS Commission III Symposium, 20-22 September 2006, Bonn, Germany.

Seidl, K., Knobbe, J., Schneider, D., Schönherr, H.-J., Lakner, H., 2006. Opto-mechanical combination of a line scanning camera and a micro laser scanner system, *Proceedings of ISPRS Commission V Symposium* 'Image Engineering and Vision Metrology', 25.-27. September 2006, Dresden, Germany.

References from websites (last accessed June 2008):
http://www.riegl.com/
http://zpr.sourceforge.net/
http://www.gnu.org/software/gsl/
http://opencvlibrary.sourceforge.net/
http://getpot.sourceforge.net/
http://www.ee.oulu.fi/~jth/calibr/
http://www.gimp.org

Graphics Applications and Visualization Techniques

DEVELOPMENT OF CAD-BASED 3D DRAWING AS A BASIC RESOURCE FOR DIGITAL RECONSTRUCTION OF BAM'S CITADEL (UNESCO WORLD HERITAGE IN DANGER)

M. R. Matini*, E. Andaroodi, A. Kitamoto, K. Ono

Research Organisation of Information and Systems, National Institute of Informatics,
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430 – (matini, elham, kitamoto, ono)@nii.ac.jp

KEY WORDS: CAD, 3D drawing, heterogeneous data, digital reconstruction, Citadel of Bam

ABSTRACT:

Digital 3D modelling of the interior and exterior spaces of destroyed architectural heritages can be done on the basis of their 2D drawings (plans, sections and elevations). In some cases like the historical Citadel of Bam (In Iran) that was destroyed in an earthquake, documentation of the site before it was destroyed is scarce and the original shapes of many buildings cannot be precisely recovered from 2D drawings made before the destruction. Instead, the essential information for 3D modelling can be derived from 2D drawings and other heterogeneous data such as photos, 3D cartographic data, geometrical similarities, and the knowledge of experts. Simultaneous application of such heterogeneous data directly by the 3D modellers leads to confusion about the domain knowledge. Other disadvantages of this process include errors in the geometry and dimensions and slow modelling. In this paper, we suggest an approach for solving these problems by developing a CAD-based 3D drawing. We model the important geometrical information of target buildings, after complementary application of heterogeneous data, as a line based (wire frame) 3D drawing in a CAD environment. This drawing can be developed by architects with average 3D drawing skills and without resorting to complex surface modelling tools. Having non-domain experts provide the complete model inside 3DCG tools like 3ds Max® is much easier and faster with our output as a basic resource. We describe the process of preparing 3D drawings and its application to 3D modelling. We indicate the advantages of our approach for precision 3D reconstruction of lost architectural heritages.

1. INTRODUCTION

The Citadel of Bam was a precious Middle Eastern historic city with adobe architecture. The city grew up from the silk and cotton trade along the Silk Roads (Mehriar, 2003). It was destroyed in an earthquake in 2003. Reconstruction of the citadel covering 180,000 square meters area will be a great challenge. Our proposed solution before physical reconstruction is a digital reconstruction of the citadel. We have been making a three dimensional model for a virtual reality realization of the Citadel as part of the Digital Silk Roads (DSR) Project since 2005 (Ono, 2005). Our work consists of several parts. The main job is the 3D modelling of buildings. The large number of buildings, different levels of importance of buildings, and availability of appropriate documents for 3D modelling are reasons for dividing the citadel area into three different levels of importance for the modelling.

1.1 Low Importance Areas

These areas consist mainly of the residential buildings and are the largest of the three areas. A variety of environmental conditions had left these buildings in ruins and no renovations had been done on them before the earthquake. The main documentation on them is aerial photos. The 3D models of exterior surfaces of these buildings can be made automatically by inspecting the lengths and depths of building shadows in the aerial photos.

1.2 Moderate Importance Areas

The surrounding walls and some gates of the Citadel make up these areas. There are several available photos of these constructions from before the earthquake, taken from different angles. The photos aren't of professional quality and were not taken with calibrated cameras. But aerial photos have better options to be used for this purpose. These resources can be interpreted through a semi-automatic photogrammetric method for moderately precise 3D modelling of exterior surfaces of the buildings (El-Hakim, 2005; El-Hakim, 2006).

1.3 High Importance Areas

The most important buildings had been almost completely renovated before the earthquake, and they have a variety of available documents. Destruction of these buildings after the earthquake makes laser scanning impossible, besides photogrammetric method cannot produce a precise model both from interior and exterior of the buildings due to lack of photos, their low resolution and lack of calibration.

This paper is about 3D modelling of these areas. The 3D models of buildings in this area are made manually. The complexity of free-form surfaces of these adobe constructions is the reason that 3ds Max® is directly used to make the 3D models (Ono, 2008; Andaroodi, 2007). Lack of precise 2D technical drawings and difficulty the 3D modellers have in applying heterogeneous data made it necessary to develop a new technical drawing. This new drawing had to be developed on the basis of heterogeneous data, and it had to be easy for the 3D modellers to use. Figure 1 divides the citadel area into three different levels of importance for the modelling; (Reference of original photo: Digital Globe), gray: low importance areas, yellow: moderate importance areas, red: high importance areas.

The best solution to the problem is a CAD -based 3D drawing. In the following section, we explain the heterogeneous data and

* Corresponding author

3D modellers' difficulties in using these data. After that, we describe the process of developing a 3D line based drawing in a CAD environment on the basis of heterogeneous data. In the last section, we evaluate the advantages and disadvantages of these 3D drawings for the process of digital reconstruction of the important buildings of the Citadel.

Figure 1: Citadel of Bam and three levels of importance

2. MANUAL 3D MODELING

Manual 3D modelling is developed usually based on 2D drawing of buildings. 3D modelling is sometimes made completely in CAD environment and imported to other appropriate software like 3ds Max® for texture mapping, lighting, rendering and animation. In case the shapes are sophisticated the 3D model is made partially in CAD or completely in other environments that have better options for modelling complicated volumes and freeform surfaces. For new buildings that are in the design process, 3D models are made on the basis of 2D drawings, sketches of architect and physical models. 3D models of an existing building (new or historic) are prepared on the basis of available 2D drawings and complementary drawings that are made directly from the building. In case a building or a complex has been destroyed, 3D modelling is made through available drawings and any other data that can help for understanding the geometry, dimensions and details. These are heterogeneous data that have different kinds of information about the buildings.

In 3D modelling of Bam Citadel, the major parts of buildings are destroyed by earthquake and the models must be developed on the basis of heterogeneous data. For this reason we start our discussion by introducing these data.

3. HETEROGENEOUS DATA

Different kinds of data formed the main resources for 3D modelling of Bam Citadel. Each kind of data consisted of appropriate 3D information about the buildings. Sometimes, a 3D modeller finds such data difficult to use by his/herself and needs to consult an architectural expert for guidance. A complete understanding of each building's form is necessary to combine different heterogeneous data sources (Andaroodi, 2007b). Below, we describe these data.

3.1 2D Drawings

Two-dimensional technical drawings, like plans, elevations and sections of different buildings of the Citadel, consisted of drawings surveyed before and after the earthquake.

3.1.1 Drawings surveyed before earthquake: these drawings were surveyed by the experts of the Bam restoration project (Iranian Cultural Heritage, Handicrafts and Tourism Organization (ICHHTO)) before the earthquake. They aren't reliable resources for 3D modelling, because there is dimensional incompatibility between the different drawings and errors in their coincidence with the 3D cartography. Despite these problems, they do show the main geometrical information about the interior and exterior spaces of the buildings.

3.1.2 Drawings surveyed after earthquake: the drawings that were surveyed before earthquake were modified and completed after the earthquake by the Bam recovery project office (ICHHTO). These drawings are more reliable than the previous drawings, but because of the destruction wrought during the earthquake and small chance of recovering their original form, new drawings could not be made to illustrate all parts of each building with an equal level of precision.

3.2 Photos

Photos of Citadel of Bam consist of on-site photos, aerial photos, and snapshots taken from movies.

3.2.1 On-Site photos: these photos consist of ones taken by ICHHTO during restoration, special events, and ceremonies, and ones taken by tourists. They are mostly of low resolution and often show popular views. Unfortunately, there are few photos of the interior spaces and the less important buildings (e.g. residential districts). There are no data related to the settings of the cameras and no photos from calibrated cameras. These photos are a major complementary resource for 3D modelling; they show different periods in the restoration undertaken within the past decade during which time the buildings undergoing restoration changed their shapes; therefore, the use of these photos had to be prioritised and they had to be annotated by architectural experts. The state of the buildings before the earthquake (after the last renovation process) served as the main guide for the 3D modelling.

3.2.2 Aerial photos: these photos were taken at seven different times during 50 years. Their scales were more than 1:2000, and they were provided by the Iranian National Cartography Centre (NCC) and Digital Globe.

Although the scale of the aerial photos is too small for them to be used directly as a resource for 3D modelling of the Citadel, they are important as a complementary data for modelling. Some of these photos aren't perfectly vertical and some building facades are visible in them. In some cases, there were no 2D drawings or normal photos available, so aerial photos had to be used as the main resource.

3.2.3 Photos from movies: still photos taken from movies are resources for 3D modelling of high importance parts of the Citadel. The movies were taken by professionals or by amateurs. They were shot from normal or helicopter views. They are valuable references for spaces and buildings for which no 2D drawing or any other photo is available.

3.3 3D Cartography Map

This 3D model was reconstituted from aerial photos taken in 1994. This photogrammetric material was made available under the Irano-French 3D Cartographic Agreement on Bam (IFCA) and the Iranian National Cartographic Centre (NCC). The cartography map was developed on the Micro Station tool and imported as an AutoCAD file with a DWG extension. The map is a wire frame 3D model consisting of lines, not surfaces or solid volumes. This 3D map makes it possible to get the XYZ coordinates of any element of the buildings in the UTM.

The 3D cartography map is useful for checking the correctness of drawings made before the earthquake and for development of 3D models of free-form mud brick walls and curved roofs. Although these maps are the most reliable reference for checking the 3D models, they have certain shortcomings: the cartography map is precise to within 5 cm in relative value and to within about decimetre in absolute terms. Therefore, its lines need to be adjusted before it can be used for 3D modelling. Moreover, the map lacks details of the facades.

3.4 Expert Knowledge

Different forms of expression comprised the expert knowledge.

3.4.1 Oral explanations: the oral explanations of experts who worked at the citadel before the earthquake are very important when no other information can be found on the geometry and structure of the buildings.

3.4.2 Texts: texts are used for disambiguating the changes in shape that occurred during the restoration and for describing the spaces. The textual references include annual reports on the restorations, unpublished papers, papers submitted to the first and second congress of the history of Iran and the Citadel of Bam and the Bam region report.

3.4.3 Sketches: sketches that are prepared by noting geometrical similarities between some parts of buildings and also on the basis of symmetry of plans can demonstrate the form and details of components of buildings that do not have any other information resource.

4. DEVELOPING CAD-BASED 3D DRAWING

The first phase of developing the 3D models from 2005 to 2007 revealed problems related to simultaneous use of heterogeneous data by the 3D modellers (Ono, 2008; Andaroodi, 2007a). The modelling process was very slow, and the models had several errors that after determination by architectural experts had to be modified or remodelled.

To overcome these problems, we prepared unified and precise data combining all the heterogeneous data and information on the geometrical forms and details of the spaces and buildings. In particular, we developed a CAD-Based 3D technical drawing (a wire frame model) as a unified basic drawing of the Citadel. This model was prepared by complementary application of heterogeneous data, and it can easily be imported to a 3ds Max® environment and completed by 3D modellers who have little domain knowledge.

There are many styles of traditional architecture whose buildings, thanks to their materials (e.g. wood, stone or bricks), have simple definable geometries for which 2D drawings are enough to show the whole form of buildings, but for this project, the 3D drawings were useful because they directly reflected the special character of the adobe buildings of the Citadel of Bam.

The important point is that architectural experts develop the 3D wire frame drawing; it is they who can comprehend the content and chronology of data and the original shapes of mud brick buildings. The drawing was developed in a sequential modification process. Below, we discuss the different stages of this process.

4.1 Initial 3D drawing

The first step is the initial 3D CAD drawing. After preparing this initial drawing, we modified it by referring to the heterogeneous data. When there were different available data, we had to choose the most appropriate one for developing the initial drawing. The importance of a piece of data is related to its quality for showing the geometrical characteristics of a space or a building. The processes of developing initial 3D drawings for interior and exterior spaces are not similar to each other.

4.1.1 Interior spaces: the initial drawing for an interior space is developed from 2D drawings. In some cases, we used 2D drawings surveyed after the earthquake, but in many cases only the 2D drawings surveyed before earthquake were available. If we have enough data for the whole space, the 3D initial drawing can show every geometrical element; otherwise, the geometrical characteristics must be completed in the next steps. Figure 2 shows the initial 3D drawing of a Sistani house's room in Citadel on the basis of its plan (Reference of plan: © ICHHTO). This plan is the only available 2D drawing of the interior space therefore form of ceiling, height of niches and other vertical elements cannot be modelled without use of other resources.

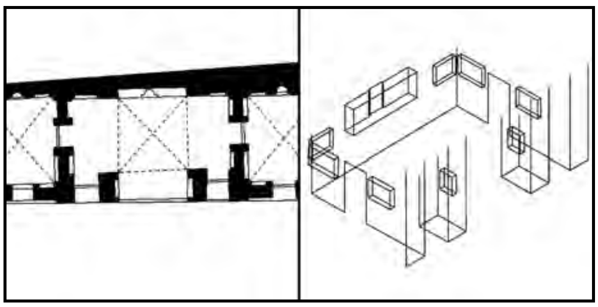


Figure 2: Initial 3D drawing of a Sistani house's room

4.1.2 Exterior spaces: 3D cartography can be directly used as an initial drawing for exterior spaces and buildings. The 3D cartography has more precise XYZ coordinates of the buildings in comparison with the 2D drawings. This 3D initial model is very important for showing the whole form of a large building or broad space; its shortcoming is the lack of geometrical characteristics of the facades. Figure 3 shows on the left: 3D cartography of the first part of the Bazaar in Citadel (Reference of cartography: IFCA (©CNRS & NCC)) and on the right: initial 3D drawing. For preparing this 3D drawing, unnecessary parts of the cartography have been deleted and the layers have been changed. Because of restrictions of 3D cartography, there are several missing lines of walls and roofs in the initial 3D drawing.

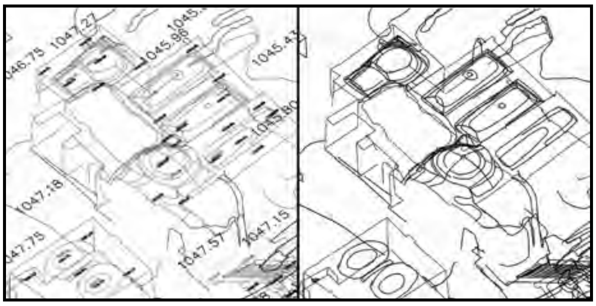


Figure 3: Initial 3D drawing of Bazaar

4.2 Modifications

Initial 3D drawings are adjusted and completed in different modification steps. In each step, one architectural aspect of the model is evaluated and modified through application of one or more heterogeneous data.

These architectural aspects are: geometry, structure, proportion, and details. The modification consists of four steps. The modification is sequential, not simultaneous, in order to avoid confusion and missing important aspects in the modification process.

4.2.1 Geometrical modification: The geometrical aspects of the initial model are evaluated in order to identify incomplete parts that can be completed with one or more appropriate data. The initial 3D drawing developed from the 2D drawings of the interior spaces is completed by referring to on-site photos and still photos from movies. The initial 3D drawing of the exterior spaces in the 3D cartography map is completed by referring to 2D drawings, on-site and aerial photos, and movie photos.

This stage is similar to a detective's investigation. There are usually several documents that don't have a direct relationship with the missing spaces, but sometimes a small sign can be a great help for completing a model. For example, sometimes a shadow on the floor or wall can reveal the existence of a small space or an architectural element that does not have any photo. Sometimes a roof in an aerial photo can be of help to determine the positions of interior walls under the roof. Sometimes, the similarity between different spaces or different elements of one space can be of help to revealing the form of the whole space.

This modification step results in a completed wire frame 3D drawing that shows all edges and corners of volumes that form an interior or exterior space (like floors, walls, roofs, domes, niches and etc.). Note that we have at this stage only the general form of a building or space without ornaments or elements like doors or windows. Figure 4 demonstrates the process of completing the initial model of a Sistani house's room through usage of heterogeneous data (on-site photos, expert knowledge, sketches), up-left A, B are the only available photos of interior space of room (Reference of photos: ©ICHHTO); C, D are similar roof construction in Bam (through guidance of experts), (Reference of photos: ©Elham Andaroodi), up-right are sketches through guidance of experts, bottom-left is initial 3D drawing, bottom-right is 3D drawing after geometrical modification.

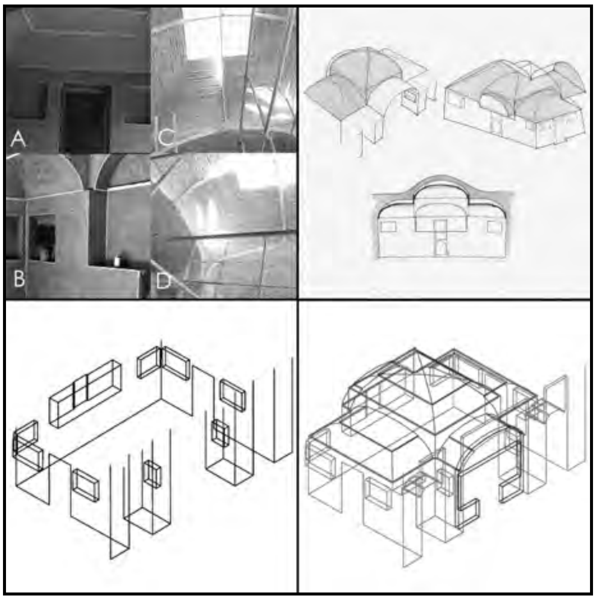


Figure 4: Completing the initial model

We use a layer management to simplify the usage of the 3D drawing and classify its large number of lines in different positions in space. We developed a layer naming system to separate the different construction elements (e.g. interior and exterior parts of roofs, walls, floors, doors and windows). This layer management enables a 3D modeller deal with the complexity of the 3D wire frame model, and it can be used to add semantic content to each component of the model. Systematic naming of layers enables us to easily show different construction elements of buildings independently for analytical research. The next steps are for adjusting and correcting the completed 3D drawing.

4.2.2 Structural modification: The structure of different parts of a building are evaluated and modified by referring to expert knowledge. These structural evaluations consist of the following steps.

The first part aims to get the correct form of domes and arches. Traditional Persian architecture has many different domes and arches, and these vary according to the period of construction, region, style, and construction materials. Adobe structural elements must be controlled, and the models must be able to be modified. This modification step relies on advice from experts and on-site photos. Usually a dome is a rotational form that in a 3D drawing has a round base and a spline as a guide for the curvature of the dome. Overall, the domes in the Citadel of Bam aren't symmetric and for this reason, we use 3D drawings of the base and two orthogonal splines for showing the domes. Figure 5 shows geometrical and structural modification of roofs and walls of the first part of the Bazaar through usage of on-site and aerial photos. up-left is on-site photo (Reference: ©ICHHTO), up-right is aerial photo (Reference: ©NCC), bottom-left is initial 3D drawing, bottom-right is 3D drawing after geometrical and structural modification

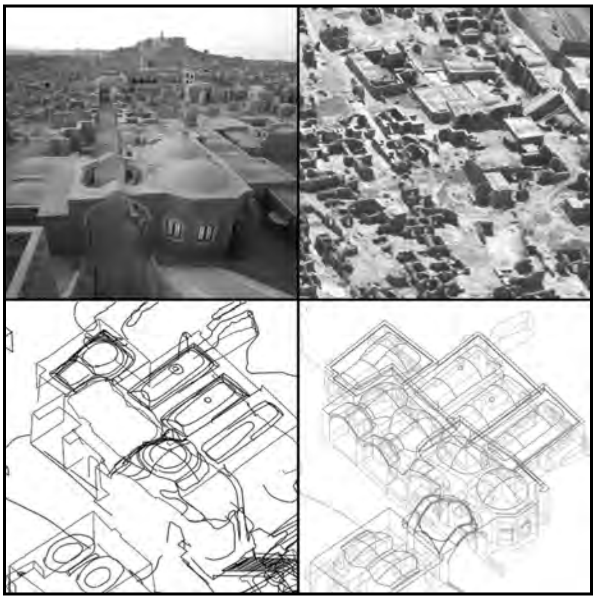


Figure 5: Geometrical and structural modification

The other structural evaluation is to control the relationship between interior and exterior elements when they are divided by the surrounding walls of buildings. The complexity of the arch and dome structures makes this kind of control necessary. Such control of a 3D model made of surfaces and solids isn't easy because we can see neither the interior nor the exterior parts of the structural elements and walls can interfere with our view. But this control and modification is much easier in a wire frame 3D model.

The correct form of a structure in some cases can be identified with the help of on-site photos taken after the earthquake. Sometimes these photos show a broken dome or wall. In these photos, we can clearly infer the construction method, the layers that make up these elements, and other useful information about the structure.

A wire frame 3D model cannot show the form of a space with surfaces, light, and shadows, but it does show how different elements of the structure are related and combine with each other.

4.2.3 Proportional modification: this step is related to control and modification of the thickness, height, angle and other dimensions of the buildings and spaces.

The curvature of outside surfaces at a domes' edge is not similar to that of the inside ceiling; hence, the roof thickness (distance between ceiling and outside surface) should be able to be controlled and modified. A 3D drawing of a dome with a base and two orthogonal splines for the outer and inner surfaces enables us to compare splines and thereby draw a roof with the correct proportions.

This step is also used to control and modify the niches, columns, and other architectural elements. By calibration of on-site photos, the proportions of these building elements can be modified. The photos are calibrated by finding the proportions of a target part in comparison with the whole building or some scalable component. By using the perspective and vanishing points in each photo, we can geometrically compare the proportions of different elements and make 3D drawings with the correct proportions. Figure 6 shows proportional modification of one part of façade on the basis of one calibrated photo of Bazaar

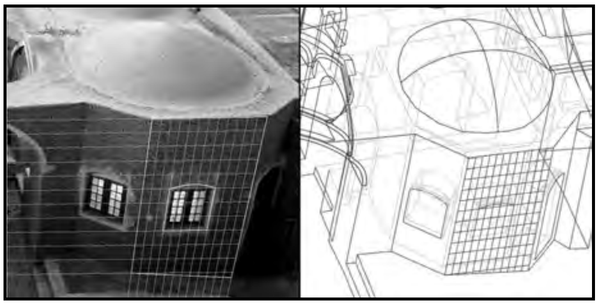


Figure 6: Proportional modification of one part of façade

4.2.4 Details: there are different architectural details that can be controlled or drawn in this step through usage of on-site photos. The combination of different architectural elements (e.g., walls and roofs meeting at the edge of roofs) is controlled in this step. Doors, windows, special ornaments, etc., are drawn in this step.

Special details like chalk bands and ornaments like muqarnas must be precisely drawn. Because of their complexity, 3D modelling of these elements without a precise 3D drawing will likely be problematic. The layer management system gives different layer names for ornaments, doors, windows, etc. Figure 7 shows completion of the details (fireplace and chalk bands) of 3D drawing on the basis of one photo taken after earthquake

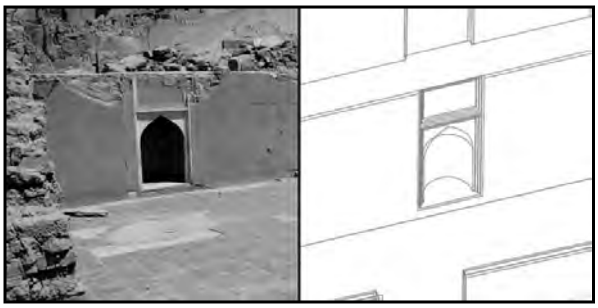


Figure 7: Completing the details (fireplace and chalk bands)

5. APPLICATIONS AND PROPOSED IMPLEMENTATION

Here, we discuss the advantages and disadvantages of our method and the results of the case study. We conclude with an example implementation of CAD-based 3D drawing for completing the 3D model of the Citadel of Bam.

5.1 Disadvantages

The following are some of the disadvantages of preparing and using CAD-based 3D drawings instead of heterogeneous data.

5.1.1 Preparation: modelling the 3D wire frame of buildings and spaces of the Citadel in a CAD environment faced several difficulties similar to 3D modelling in 3ds Max®. A unified preliminary and precise 2D drawing is not available for 3D modelling, and the semantics of heterogeneous data are complicated. Hence, the 3D drawing must be made by someone that is not only an architectural expert but also familiar with 3D CAD drafting. Even under supervision of an architect, a CAD operator will duplicate tasks and waste time in such an effort. Model components are developed by complementary comparison of different data; hence, the person drawing it is the one who needs to comprehend it.

5.1.2 Time: although 3D modelling in Max on the basis of a 3D CAD drawing is much faster and subject to fewer errors, we must not forgot that preparing a precise and correct 3D drawing needs a considerable amount of time for preparing the initial drawing and for the four modification steps. In other words, the complexity of 3D modelling forces us to have the extra process of creating a CAD-based 3D drawing.

5.1.3 Restrictions: 3D drawings are good for modelling flat surfaces or one curved surface but they are not good for modelling two or more curved free-form surfaces. The Citadel of Bam consists of adobe constructions in the desert Middle Eastern style of architecture, and for this reason, the roofs, walls, and other construction elements have many curved surfaces. We need a special way to depict these surfaces as lines and also guidelines for modellers to choose the appropriate way to convert these lines into surfaces and complete the model.

5.2 Advantages

3D drawings have several advantages for 3D modelling of the important buildings of the Citadel.

5.2.1 Advantages over 2D drawings: a 2D drawing (e.g. plan or section) can show horizontal or vertical sides of a building with XY, YZ, or XZ coordinates whereas the 3D drawings contain all vertical and horizontal images with XYZ coordinates for every element of the buildings. The spaces with rectangular edges can be easily shown in 2D drawings, but many buildings in the Citadel of Bam don't have rectangular edges; there are many curve surfaces that cannot easily be shown in these drawings. The problem of using 2D drawings in modelling the upper part of the Citadel is critical. The upper part is the governor quarters, and it was built on a hill. There are many floors and roofs that are not on the same level and spaces on different levels that are combined irregularly together because of the special topography of the ground.

For this reason, the 2D drawings drawn before the earthquake have several errors, and a comparison with 3D cartography clearly shows their dimension problems. A 3D drawing is a much better resource for reconstruction of this quarter.

5.2.2 Unified basic data: the most important reason for using 3D drawings instead of heterogeneous data is that it relieves the 3D modeller. Unified data is a bridge between the different available data and 3D modellers. Although preparing a 3D drawing isn't fast and is complicated, our experience in the first phase of 3D modelling (without 3D drawings – between 2005 and 2007) shows that a modelling route whereby architectural experts variously evaluate the 3D models and CG experts then modify and correct the models takes more time and is much more difficult.

Without this 3D drawing, the modeller must start with 2D drawings containing errors and missing data. The 3D cartography can also be used, but its large scale means it can't be used to make a precise 3D model with a high level of detail. To modify the 2D drawing and 3D cartography, the modellers need to be able to simultaneously use other data. They aren't specialists in traditional architecture, and they cannot determine which data is more reliable than the other. Therefore, these complexities and ambiguities are better solved by architectural experts before the modeller starts his/her work.

5.2.3 Independence of domain experts and modellers: the first phase of 3D reconstruction in this project was preparation of appropriate heterogeneous data by architectural experts and their delivery to 3D modellers. In the first phase, the tasks of the experts included control of the 3D models, reports of errors, and control of the process of modification or remodelling by the modeller. This situation needed a constant relationship between the modellers and architects, and the domain experts were overloaded with tasks. A 3D drawing made by architectural experts decreases the dependency of modellers on domain experts, reduces the task overload of both sides, and leads to fewer remodellings and modifications.

5.3 Implementation by modeller

From 2005 to 2007 that the first phase of 3D modelling of seven buildings was almost completed, application of heterogeneous data by 3D modellers caused several problems. We evaluated 3D model of each building during development at least 3 times and specified each time between 100 to 200 errors (architectural and technical) to be modified. We provided evaluation reports to request corrections and also organized meeting and discussions with modellers. However the correction process was complicated and even remodelling was easier than modification for some errors.

For solving this problem in second phase of the project (since October 2007) we provided a 3D drawing developed in a CAD environment for the 3D modellers of our project. The modellers were students of architecture who have expertise in 3ds Max® (from Espace Virtuel de conception en Architecture et Urbanisme (EVCAU), Ecole Nationale Supérieure d'Architecture Paris-Val de Seine (ENSAPVS)), students of computer graphics without knowledge of architecture (from Global Information and Telecommunication Institute (GITI), Waseda University, Tokyo), and 3D modellers (Raazahang, University of Tehran (UT)).

They imported the DWG file to 3ds Max®. Their task was only to define faces between the borders of 3D drawing lines. This reduced the synthesis phase of original shapes by them specially for arches, vaults or domes and proportions of niches. In most cases, the lines were given for every detail and the only challenge was to choose proper surface modelling or modifier among large options of the tool, for example polygons or

NURBS* for cloister vaults and mesh for barrel vaults (Ono, 2008).

3D model of "Bazaar" is an example for comparison of our previous approach (first Phase) and new approach (second phase). This model was the most problematic one among other seven buildings of the first phase due to different variations in its long form. Successive vaulted passageway had also specific traditional geometry with some non renovated parts. The modelling had several errors and 3 phases of evaluation could not solve the problems. We decided to completely remodel it. In second phase the modeller applied our 3D drawing and completed it in less than three mounts with minor errors to be corrected. Figure 8 shows the implementation of interior and exterior 3D drawings for developing the final 3D model (one Sistani room and one part of Bazaar),

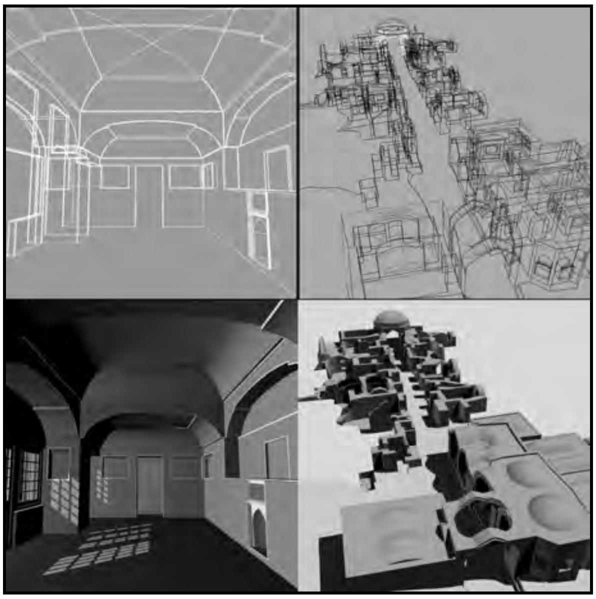


Figure 8: Implementation of interior and exterior 3D drawings

5.4 Conclusion

In this paper, we explained the development of a CAD-based 3D drawing for a 3D digital reconstruction of the UNESCO world heritage in danger, the Citadel of Bam. We described the importance of CAD-based 3D drawing for our case study that deals with buildings destroyed in an earthquake, a lack of precise preliminary 2D drawings, a diversity of complementary resources for modelling, and complex forms of adobe buildings with traditional arches, vaults, and details. We used complementary heterogeneous data in the four stages of the modifications in order to prepare the 3D CAD drawing. The modellers used the 3D drawing as a unified basic material for 3D modelling without consulting domain experts.

This approach is a key process for large-scale and high-precision manual 3D modelling especially when no building is left to be laser scanned or photo modelled. Our method is a solution for problems around the globe where similar heritage sites are damaged or in ruins as result of natural or manmade disasters and where there is little documentation on the original shapes of architectural artefacts.

* Non-uniform rational B-spline

REFERENCES

Andaroodi, E., Matini, M. R., Abe, N., 2007a. 3-D Reconstitution and virtual reality of the world heritage in danger, Citadel of Bam. Proceeding of short papers and posters *IPT- EGVE, 13th Eurographics Symposium on Virtual Environments*, Editor: Dieter Fellner, Weimar, Germany, pp. 43-44

Andaroodi, E., Matini, M. R., Abe, N., Ono, K., Kitamoto, A., Kawai, T., Mokhtari, E., 2007b. Simultaneous implementation of heterogeneous data for 3-D reconstitution of the UNESCO world heritage in danger: Arg-e-Bam. Proceeding of the *Jinmonkon*, *IPSJ symposium series*, Kyoto, Japan, Vol. 2007, No. 15, pp. 265-270

Bam, Citadel of Bam, heritage in danger, Digital silk road website, National Institute of Informatics, http://dsr.nii.ac.jp/bam/ (accessed 22 August. 2008)

Chopin, F., Bouet, O., Andaroodi, E., Ono, K., Mokthari, E., Gholipour, V., 2007. Modélisation par objets réutilisables: fragmentation en éléments d'une architecture de terre. Application à la modélisation du petit caravansérail de la citadelle de BAM - 4p - Proceedings of *Virtual Retrospect 2007* , Université de Bordeaux 3, France

Dosch, P., Masini, G., 1999. Reconstruction of the 3D Structure of a Building from the 2D Drawings of its Floors, *Fifth International Conference on Document Analysis and Recognition (ICDAR'99)*

Dudek I., Blaise J.Y., 2003. Access, interpretation and visualization of heritage data using the architectural morphology: experimenting emerging interfaces on a case study. ICHIM 03- Archeology & Architecture

El-Hakim, S. F., Beraldin, J. A., Gonzo, L., Whiting, E., Jemtrud, M., Valzano, V., 2005. A Hierarchical 3D Reconstruction Approach for Documenting Complex Heritage Sites*, The XX CIPA International Symposium.* Torino, Italy. NRC Publication Number: NRC 48229

El-Hakim, S. F., MacDonald, G., Lapointe, J. F., Gonzo, L., Jemtrud, M., 2006. On the Digital Reconstruction and Interactive Presentation of Heritage Sites through Time, *7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2006),* Nicosia, Cyprus, NRC 48782

Georgopoulos, A., Ioannidis, C., Daskalopoulos, A., Demiridi, E., 2004. 3D reconstruction and rendering for a virtual visit. In: *XX ISPRS Congress,* Commission V, Istanbul, Turkey, IAPRSSIS Vol. XXXV, Part B5, pp. 632-636.

Gonzo, L., Voltolini, F., Girardi, S., Rizzi, A., Remondino, F., El-Hakim, S. F., 2007. Multiple Techniques Approach to the 3D Virtual Reconstruction of Cultural Heritage, *Eurographics Italian Chapter Conference 2007,* R. De Amicis and G. Conti (Editors)

Mehriar, M., 2003. The history of Citadel of Bam, (in Persian), Bam recovery office, Iranian Cultural Heritage and Tourism Organization, Tehran, Iran

Ono, K., Andaroodi, E., Einifar, A., Abe, N., Matini, M. R., Bouet, O., Chopin, F., Kawai, T., Kitamoto, A., Ito, A., Mokhtari, E., Einifar, S., Beheshti, M., Adle, C., 2008. 3DCG reconstitution and virtual reality of UNESCO world heritage in danger: the Citadel of Bam. *Progress in Informatics*, National Institute of Informatics, Tokyo, No. 5, pp. 99–136.

Ono, K., Yamamoto, T., Kamiuchi, T., Kitamoto, A., Andres, F., Sato, S., Andaroodi, E., 2005. Progress of the Digital Silk Roads project. *Progress in Informatics*, National Institute of Informatics, Tokyo, No. 1, pp. 93-141

ACKNOWLEDGEMENT

3DCG reconstitution and virtual reality of the citadel of Bam is a collaborative research project between ICHHTO and DSR project of National Institute of Informatics. ICHHTO provided part of resources that are introduced in this paper and local support of surveys or discussions in the Bam site.

CURATION AND PRESERVATION OF CAD ENGINEERING MODELS IN PRODUCT LIFECYCLE MANAGEMENT

M. Patel [a], A. Ball[a], L. Ding[b],

[a]UKOLN, University of Bath, Bath, BA2 7AY UK -(m.patel, a.ball)@ukoln.ac.uk
[b]IMRC, Dept. Mechanical Engineering, University of Bath, Bath, BA2 7AY UK -ld218@bath.ac.uk

KEY WORDS: Curation, Preservation, Engineering, PLM, CAD Models, Multilayer Annotation, Representation Information

ABSTRACT:

During the last decade the management of a product's data over its entire life has been gaining prominence. This is largely due to two factors: firstly, an increasingly collaborative environment in which product development and maintenance takes place in a geographically distributed and networked environment. Secondly, there is an emerging economic paradigm shift in which companies that design and build products are increasingly entering into contracts to provide through-life support for them - that is, products are being purchased as services rather than artefacts. For engineering companies, this shift entails a commitment to supporting products over a much longer timeframe than previously expected. At the same time, there is an increasingly greater reliance on CAD models which are now being used as the method for recording definitive product data. However, the CAD software industry is characterised by ephemeral, backwardly incompatible, proprietary applications and file formats which readily become obsolete, making the long-term retention and accessibility of digital product models and data a challenge.
We examine the curation and preservation requirements in Product Lifecycle Management (PLM) and suggest ways of alleviating the problems associated with the sustained representation of CAD engineering models through the use of lightweight formats, layered annotation and the collection of *Representation Information* as defined in the Open Archival Information System (OAIS) Reference Model.

1. INTRODUCTION

The emergence of Product Lifecycle Management (PLM) as a business model over the last decade or so can be attributed to the disruptive effects of ICT which has resulted in a global and networked market place and consequent international collaboration and business practices. This business model is applicable in the engineering, manufacturing, contracting and service sectors amongst others. PLM requires the efficient capture, representation, organisation, retrieval and reuse of product data over its entire life [McMahon et al., 2005].

The importance of managing a product's data over its entire lifecycle is gaining importance, mainly due to two factors. Firstly, many companies now operate in an increasingly collaborative environment, one in which product development and maintenance occur in a geographically distributed and networked environment, with the result that much of the data relating to a particular product or artefact is dispersed over a number of organisations and locations. Secondly, there is an emerging economic paradigm shift such that companies that design and manufacture products are increasingly entering into contracts to provide through-life support for them – that is products are now being sold and purchased as services rather than artefacts. For example, within the aerospace industry, Rolls-Royce has introduced the concept of "power by the hour". For products such as cruise ships, aircraft, rolling stock for railways, hospitals and schools, this could mean a commitment to providing support for as long as the product is in service, extending to 30-50 years or in some cases even longer.

At the same time, there is a much greater reliance on CAD models which are now being used as the main carriers for recording definitive product data as opposed to paper based technical drawings and documentation. Within the last five years or so, the engineering industry has moved over to using CAD models directly for communicating designs, not only to manufacturers and builders, but also to regulating authorities and maintenance crews. However, the switch to recording information digitally presents its own problems, not only in terms of long-term maintenance and accessibility, but also as a potential threat to the recording of the evolution of design, artefacts and products in terms of our industrial, automotive and avionic history and heritage.

The remainder of this paper examines curation and preservation issues in the context of PLM and suggests ways of alleviating the problems associated with the sustained representation of CAD engineering models through the use of lightweight formats, layered annotation and the collection of *Representation Information* as defined in the Open Archival Information System (OAIS) Reference Model [CCSDS, 2003].

2. DIGITAL CURATION AND PLM

Digital Curation

The term *digital curation* is now generally accepted as including the active management of digital data over their useful lifetime, both for contemporary and future use, as well as incorporating archiving and digital preservation. The term also encompasses the notion of adding value to a trusted body of digital information as well as its reuse in the derivation of new information and the validation and reproducibility of results [Beagrie, 2006; DCC, 2007].
The urgency for assuming widespread digital curation activities stem from several issues all of which relate to the proliferation of digital information and the heavy risks associated with its potential loss. A study, conducted by UC Berkeley estimates that the world produces between one and two Exabyte of unique information every year [University of California, Berkeley, 2000]. A recent follow on report forecasts that the "digital universe" will explode to an incredible 988 Exabyte by the year 2010 [IDC White Paper, 2008]. Additionally, legislative and regulatory requirements imposed on certain industries such as

pharmaceuticals and engineering mean that they are required to maintain data and records for considerable periods of time.

The current situation is such that while the cost and investment in digital information creation is huge, the benefits are likely to be short-lived and the threat of the “Digital Dark Ages” [Kuny, 2007] will remain omnipresent unless digital information and data is curated and preserved adequately. Due to technological obsolescence (hardware, software and file formats), which constitutes one of the major threats to digital information, data can become inaccessible within a very short time. Moreover, much digital information requires software applications in order to make is accessible to humans.

A variety of techniques have been proposed and explored to combat the effects of rapidly changing technologies and media degradation: bit-stream copying, refreshing, the use of durable media, digital archaeology and replication. Strategies aimed at preserving access to the information content and providing functional preservation include: technology preservation, analogue backups, migration, normalization, emulation and encapsulation. A particular strategy concerned with mitigating the effects of technology evolution is based on the use of *Representation Information* (RI) – a concept used in the Reference Model for an Open Archival Information System (OAIS) [CCSDS, 2003]. RI is all-encompassing; it is essentially any information that is required to render, process, visualize and interpret data and includes: file formats, software, algorithms and standards as well as semantic information.

The Open Archival Information System (OAIS)

The OAIS Reference Model establishes a common framework of terms and concepts for use in the preservation of information. An archival information system consists of an organisation of people and systems, which has accepted the responsibility to preserve information and make it available for a *Designated Community*. The latter being an identified group of potential stakeholders and users. The Model is set in the context of producers (who generate information to be archived), consumers (who retrieve that information) and management (the wider organisation responsible for maintaining the OAIS).

The model has achieved widespread adoption, influencing: the development of preservation planning [Strodl et al, 2007]; preservation metadata [PREMIS, 2008]; architectures and systems of repositories [Giaretta, 2007]; and conformance and certification criteria for archives [TRAC, 2007]. Of particular note is OAIS PDI or *Preservation Description Information*, comprising several types of metadata to help ensure the quality of the data and its fitness for purpose:

- *Reference*: One or more mechanisms used to provide identifiers for unambiguous access to content e.g. object identifier or a persistent identifier.
- *Provenance*: Documents the history of the content information, to provide some assurance as to its likely reliability.
- *Context*: Documents the relationships of the content information to its environment and other content information e.g. calibration history; relationship to other data; or pointers to related documents.
- *Fixity*: Provides data integrity checks including validation and verification keys used to ensure authenticity e.g. encoding and error detection schemes such as checksums.

The OAIS is in effect a comprehensive reference model for the development and operation of a preservation and curation environment for all types of data.

A Digital Curation Lifecycle Model: Digital curation is a multi-faceted and complex process involving social, political, organisational and financial as well as technical issues. In order to clarify these numerous aspects and the relationships between them, the Digital Curation Centre (DCC) has developed a Curation Lifecycle Model [DCC Curation Lifecycle Model, 2008], see Figure 1.

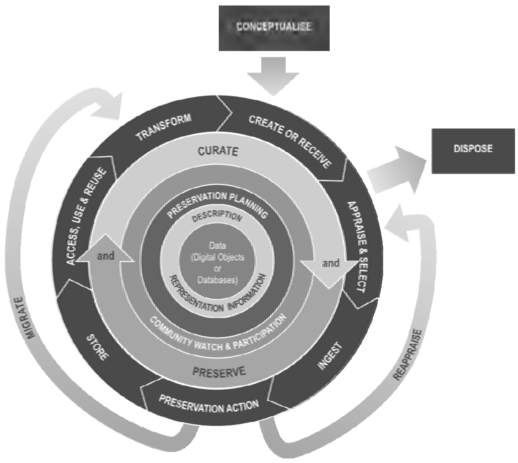


Figure 1: DCC Curation Lifecycle Model © DCC 2008

The Model provides a graphical high level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt. It can be used to plan activities within an organisation or consortium and enables granular functionality to be mapped against itself and particular information workflows. The DCC Curation Lifecycle Model defines digital objects, as well as databases and splits the processes into those that are:

- Full lifecycle stages (Description and Representation Information; Preservation Planning; Community Watch and Participation; Curate and Preserve)
- Sequential actions (Conceptualise; Create or Receive; Appraise and Select; Ingest; Preservation Action; Store; Access, Use and Reuse; Transform)
- Occasional actions (Dispose; Reappraise; Migrate)

Product Lifecycle Management

The scope of PLM is extensive and includes a number of phases: Conceptualisation (innovation, requirements); Design Organisation (people, infrastructure, knowledge); Design (product, process); Evaluation (analysis, simulation, performance, quality); Manufacture and Delivery (production, supply, delivery); Sales and Distribution (advertising, marketing); Service and Support (maintenance, upgrades, warranties); Decommissioning (retirement, recycle, disposal). The Knowledge and Information Management through Life Project (KIM) is currently investigating the implications of PLM and the paradigm shift to a product-service approach [Ball et al., 2006]. A scenario has been devised by the Project to illustrate ideal information flows in a product’s lifecycle, Figure 2 shows a summary. In all cases the flow of information must be managed to ensure that appropriate information is transmitted and understood at the different stages. A major challenge is that data needs to be shared and exchanged between multiple organisations involved in the lifecycle of the product which can extend to considerable lengths of time.

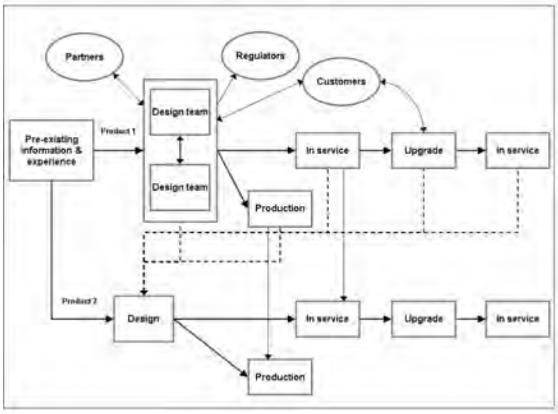


Figure 2: Information flows in the lifecycle of a product

Up until the turn of the millennium, engineering software was used to support a paper-based workflow. CAD packages were used to create virtual models of designs, from which drawings and other design documentation could be produced. The manufacture or construction process was based on the resulting documentation. However, current digital environments necessitate an electronic flow of information between heterogeneous systems for Computer Aided Design (CAD), Computer Aided Engineering (CAE) and Computer Aided Manufacture (CAM) as well as Enterprise Resource Planning (ERP), Customer Relationship Management (CRM) and Supply Chain Management (SCM). Additionally, users in the differing stages of a product's lifecycle require different information and representations - i.e. there are multiple viewpoints on the product models [Ding et al, 2006]. For example, machining features are useful for manufacturing engineers, but not for marketing staff, for whom a visualisation of a product, unencumbered with production and manufacturing information, is of far greater use.

Curatorial issues in PLM

The active management of all product data is vital to PLM as data is created, added to, modified and extracted over the course of the lifecycle of a product. It is apparent that elements of the DCC Curation Lifecycle Model (Figure 1) could be incorporated into the information flows at various stages in PLM (Figure 2). However, although current PLM systems such as Agile, IBM/Dassault, MatrixOne, PTC and UGS PLM, cater for some aspects of digital curation, they do not place an emphasis on issues relating to preservation. Additionally, the scope of PLM is wide-ranging, involving a large volume and variety of data, information and knowledge, all of which needs to be managed and maintained. This can range from highly structured data (such as geometric models and databases) to unstructured textual documents (e.g. email) to the tacit knowledge held by employees (e.g. design rationale and lessons learned from experience).

Whilst digital product information may share many issues in common with other types of data, such as text documents and scientific datasets (e.g. issues relating to bit-preservation such as information security, integrity, authenticity and longevity of digital storage media) there are several aspects which are specific to PLM and that make the digital curation of engineering product data particularly rife with problems.

Within PLM there is also a requirement to support global and distributed collaboration. However, product data tends to be amongst the most valuable intellectual property of a company, which will therefore only be prepared to share selective information depending on the role of the collaborating partner.

In addition, not all stages of PLM require all product data and it is useful to extract a simplified view of the product for use in later stages. Here the notion of *Significant Properties* comes to the fore. Significant properties are those aspects of the digital object which must be preserved over time in order for it to remain accessible and meaningful [Significant Properties Workshop, 2008]. In PLM the significant properties of a product may vary depending on the view and the stage in the lifecycle.

Due to the complexity of many contemporary products, the volume of data generated during development tends to be huge and distributed, so that it becomes difficult to make decisions regarding which data and how much should be kept and maintained.

Issues relating to technological obsolescence are exacerbated in PLM, mainly due to the state of the CAD software industry and the move to the product-service paradigm. Complex dependencies and relationships exist between file formats, software and hardware. In addition, the CAD software market is competitive and characterised by a proliferation of CAD formats which are proprietary, closed and subject to frequent change. Interoperability between such systems is virtually non-existent – indeed many CAD tools do not even maintain reliable backwards compatibility with their own versions.

Solutions such as emulation and migration pose problems in the engineering domain. Emulating old software incurs difficulties with integrating it into complex and more modern workflows and systems. A major issue with migrating old designs to newer formats is that there is always the risk of data loss and subtle design corruption. Also, the cost of re-checking and re-validating a design after migration can be substantial. Even the use of open and neutral standards, such as IGES (Initial Graphics Exchange Standard) [IGES, 1996] and STEP (Standard for Exchange Product Data) [STEP, 2005] is not without issues. The rigours and long timescales of developing a comprehensive exchange standard for CAD models, means that it is difficult to keep up to date with the latest capabilities of CAD tools. Furthermore, the level of support for such standards can be variable between tools. As a consequence, data created using a particular application is in danger of becoming inaccessible once that software is retired or replaced. Furthermore, in such a complex and dynamic environment, it becomes extremely difficult to reliably retrieve and trace provenance information to check the veracity and reliability of data. To facilitate the development of new generations of a product, especially in the face of greater awareness of environmental impact and efficiency, it is necessary to cater for long-term retention and preservation so that older designs can be reused and adapted or customised.

3. CHALLENGES FOR CAD ENGINEERING MODEL REPRESENTATIONS

CAD models have traditionally been used in the design, evaluation and manufacturing phases of PLM and product information is still largely stored in CAD models based on conventional product representations, including boundary representation (B-rep), freeform surface modelling, feature-based models or parametric models. If CAD models are to become the main carriers of extra product lifecycle information, it is clear that they need to be extended, augmented and supported in additional ways.

PLM makes several demands that need to be taken into account:

- Protection of commercially sensitive information (Intellectual Property)
- Generation of view-point specific representations to support differing processes
- Rapid sharing of information between geographically distributed applications and users (interoperability, platform and application independence, use of standards, reduced file sizes etc.)
- Support for recording feedback from downstream processes
- Long-term preservation (recording of metadata, design rationale, open formats, RI, use of standards etc.)

4. A STRATEGY TO SUPPORT CURATION IN PLM

To extend a CAD model from purely the design stage into the whole product lifecycle, a framework of lightweight representations is proposed together with a method for annotation and the use of a Registry/Repository of Representation Information (RRoRI) to support decision making.

In the proposed strategy, all users throughout a product's lifecycle can annotate the CAD model according to their specific requirements and experiences. The information is stored in a series of separate XML-based files, each of which is linked to the CAD model through a specific element (e.g. a face) using a mechanism of references. With the support of these markup files, the CAD model can be compressed into various lightweight representations according to different levels of security and viewpoints. Additionally, RI relevant to the CAD model and lightweight formats, as well as the XML schemata for markup documents, is stored in a RRoRI to aid the interpretation of the accumulated data in the longer term. Issues relating to the collection and use of engineering RI are explored in Patel & Ball [Patel & Ball, 2007].

Lightweight CAD Model Representations

Full CAD formats tend to be large, complex and proprietary, and are rarely backwards compatible. This makes them unsuitable both for long-term archiving, reference and reuse, and for distributed collaborative design work. Lightweight formats provide a potential solution. They comprise simple formats that are easier to preserve but which do not try to retain all the richness of the full CAD model. By producing files in these formats at the time of the original design, they can be validated at the same time as the full model. Their simplicity makes them easier to read back into newer software. They have smaller file sizes, simpler and more open specifications, and more affordable software support. Also, they need only contain as much information as a particular recipient needs (this is analogous to the notion of *desiccated formats* [Kunze, 2005] which retain only essential information).

There are a number of different lightweight representations in current use, each with properties and characteristics better suited to some purposes than others. In this section we introduce a number of these formats, with a particular regard to their capabilities with respect to: fidelity to the full model, metadata storage, data security, file size reduction, support for the format by software and openness.

3D XML: 3D XML [Versprille, 2005; Dassault Systèmes, 2007] is an XML-based format for describing a model's geometry, structure and visualization, and is optimized for interactivity and compactness. It can represent geometry using compact NURBS-like surface descriptions, XML polygon meshes and compact syntax polygon meshes, but does not have any additional security features. File sizes are kept down by a reference-instance mechanism (allowing the same data to be re-used several times within a model), a modification mechanism (allowing an instance or reference object to build on the properties of another reference object) and raster graphic compression. Models may be expressed by a single file or split across several` files. Native support for the format is largely restricted to Dassault Systèmes products, although free plug-ins are available for Lotus Notes and Microsoft Word, PowerPoint and Internet Explorer, as well as a free standalone viewer. The format is owned and controlled by Dassault Systèmes; the specification for the format is available cost-free to those who register.

JT Format: JT Format [UGS, 2006] is a binary format for encoding product geometry using boundary representations and wireframes, and supports additional product manufacturing information and other metadata. It does not have any built-in security features other than approximating data using tessellating polygons. File sizes are kept down using a reference-instance mechanism, zlib compression of various data elements and datatype-specific compression using algorithms such as uniform data quantization, bit length codec, Huffman codec, arithmetic codec, and Deering Normal codec. Models may be expressed by a single file or split across several files. Native support for the format is largely restricted to UGS products, although free plug-ins are available for Microsoft Word, Excel and PowerPoint, as well as a free standalone viewer. The format is owned by UGS, but the specification is freely accessible on the Web and blanket permission is given to implement it.

PLM XML: PLM XML [UGS, 2005] is a set of XML schemata for describing a model's geometry, structure, features, ownership, and visualization. It is designed to be interoperable between a number of different tools from across the lifecycle of a product. The native schemata for representing geometry can support 2D and 3D vector graphics, NURBS surfaces and features, although non-native representations can also be used or referenced in a PLM XML document. It also allows for a single logical product model to have several different geometric representations, tailored to different purposes. Metadata of several different types – mass, material, texture, product manufacturing information, dimensions and tolerances, user markup, application-specific data – can be attached to logical parts of the model or specific geometric representations. File sizes can be reduced using a reference-instance mechanism and by splitting out various sections of data into separate files (so that data not needed for a particular purpose need not be transmitted). As well as approximating and sub-setting data, PLM XML also supports mechanisms for restricting access to parts of the model data on the basis of person, organization or place. The format is used extensively by UGS products but is not widely supported otherwise. The format is owned and controlled by UGS; the XML schemata are freely accessible on the Web, but the software development kit must be purchased.

PRC: PRC [Adobe Systems, 2007b] is a binary format that promises to encode the full range of CAD geometry, along with model trees, history trees and various forms of markup. Alternative geometries (e.g. exact and tessellated) can be provided for each part; markup can be associated with entire parts or tessellations but not with items of exact geometry. Arbitrary non-PRC data can be included at various points, notably at the end of entity code. Summary data sections enable files to be accessed without being fully parsed, but the format is not suitable for streaming. File sizes are reduced through a number of mechanisms: compact mathematical encoding of geometry, a reference-instance mechanism, and gzip encoding of data sections (header sections remain uncompressed). The precision of the geometry may be reduced to provide lossy compression. Proprietary converters are available for a wide range of CAD formats, and PRC is supported as a native 3D model format within the Portable Document Format (PDF) specification from version 1.7 (corresponding to Adobe Acrobat 8.1), which adds some conservative security measures on top of the otherwise unprotected format [Adobe Systems 2007a]. The format was initially proprietary but is expected to form part of ISO 32000.

Universal 3D (U3D): Universal 3D [ECMA-363, 2007] is a binary format for encoding product geometry using sets of tessellating triangles and (from the 4th edition) NURBS surfaces. A mesh update mechanism allows meshes to be rendered progressively, providing basic streaming support. Metadata, stored as key/value pairs, may be attached to any node in the model tree. It does not have any in-built security features other than approximating the geometry. File sizes are kept down using a reference-instance mechanism and a bit compression algorithm on numeric data fields. The format is most notably supported as a native 3D model format within the PDF specification, with the 1st edition of U3D supported from version 1.6 (Adobe Acrobat 7) and the 3rd edition supported from version 1.7 (Adobe Acrobat 8.1). PDF also adds some conservative security mechanisms of its own. It was developed by the 3D Industry Forum and is published and maintained as ECMA standard 363; the specification is freely available on the Web.

X3D: X3D (ISO/IEC 19775 2004; ISO/IEC 19776 2005; ISO/IEC 19777 2006) is an improved version of Virtual Reality Markup Language (VRML); it is an XML format optimized for animation and interaction. It can represent 2D and 3D vector graphics, 3D tessellating polygon meshes, and NURBS surfaces as well as identifying bones and joints for human animation. Any node in the model tree may have metadata attached, in a format specifying a value (string or number), a metadata schema and a key. It does not have any in-built security features other than data approximation. X3D has a reference-instance mechanism and a relatively compact XML syntax, with coordinates expressed as space/comma delimited lists within attributes, rather than through a hierarchy of tags; a binary syntax is available that compresses field values according to Fast InfoSet principles, using zlib compression, quantization of floating point number arrays, integer range reduction and conversion of absolute values to relative values. Open source libraries and viewers are available for processing and rendering X3D files. X3D was developed by the Web 3D Consortium, and is published and maintained as ISO standards 19775, 19776 and 19777; these standards are freely available on the Web.

XGL/ZGL: XGL [XGL Working Group, 2006] is an XML-based encoding of the Open Graphics Library (OpenGL) application programming interface for rendering 2D and 3D computer graphics. When compressed it is known as ZGL. It uses tessellating triangles to encode geometry, and is optimized for display. It does not have any capabilities for storing metadata, nor does it have any in-built security features other than approximating the geometry. File sizes are kept small using a reference-instance mechanism and a relatively compact XML syntax, with vector coordinates expressed as comma delimited lists rather than through a hierarchy of tags. XGL is supported by Autodesk and a few smaller CAD vendors. It was developed by the XGL Working Group but no longer appears to be maintained; the specification of the format was once freely available on the Web, but now only appears in 'unofficial' locations.

Multilayered Annotation (LiMMA)

Although lightweight representations alleviate many of the challenges outlined in section 3 (collaborative exchange, customised views, security, application independence, preservation, reduced file sizes etc.) there is still an out-standing requirement – that of being able to augment the geometric model of the product with additional information from different phases of PLM. For this purpose we propose the use of annotation which allows the incorporation of varied information:

- Design rationale, context, provenance, RI etc.
- Extra information needed for a certain point of view
- Embedding of commercial security levels to restrict access to certain partners or users

Use of a machine processible language, such as the Extensible Markup Language (XML) [XML 2006], is particularly useful for collaborative ventures over the Web.

There are two methods for applying annotation, 'inline' and 'stand-off'. Inline annotation involves adding information directly into the text of a document or model, whereas the stand-off (external or reference) method allows markup information to be stored separately, being linked back to the document or model using references or pointers [TEI Standoff Markup Working Group, 2003; Thompson & McKelvie, 1997]. The latter is the more appropriate for use with CAD models for several reasons:

- It allows a 3D geometric representation of a product to be progressively expanded to include additional metadata without changing the representation method used for the geometry of the product.
- The CAD model itself need not contain all the information required for every user and purpose: context-specific information can be extracted into a number of separate files to provide multi-layered annotation which can be passed around as required, allowing the CAD model to remain smaller in size.
- It allows the same annotation to be applied to different representations of the same model, granting the annotation information some independence of the CAD format used.
- It enables downstream processes (e.g. finite element analysis and manufacturing processes) to be independent of the CAD model through the reuse of annotation.

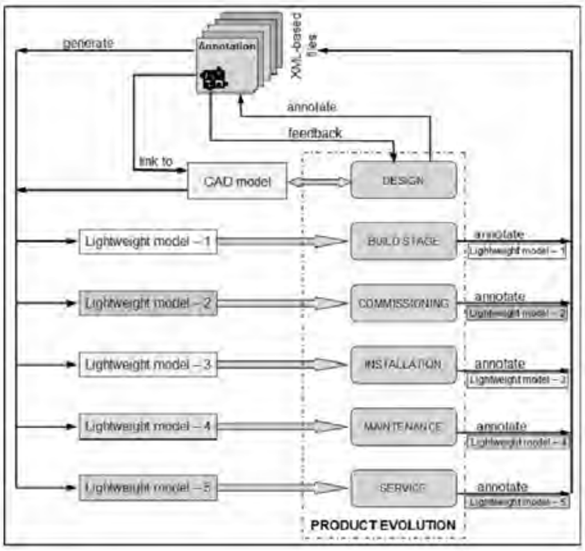


Figure 3: A Framework for the annotation of CAD models

LiMMA (Lightweight Models with Multilayered Annotations), is a framework for representing CAD models using lightweight geometric models with additional layers of XML-encoded information, as shown in Figure 3. To date, LiMMA plug-ins have been written in C/C++ and NX Open for UGS's NX CAD package, as well as in JavaScript for the 3D PDF viewer Adobe Acrobat Reader. In addition, a standalone X3D viewer has been written in Java as another component in LiMMA. Annotations are currently linked to the geometric models by means of unique identifiers attached to the entities that comprise the model. However, an alternative system of referencing making use of co-ordinate sets is also under development. Figure 4 shows the annotation environment which has been developed within the NX package.

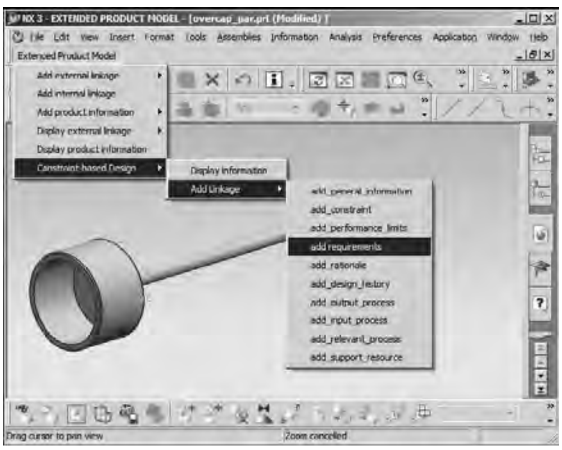


Figure 4: Interface of the internal NX markup environment

Representation Information (RRoRIfE)

Although the use of lightweight formats and multi-layer annotation seemingly address all of the challenges outlined in section 3, there is a further issue with the proposed strategy in that a wide range of lightweight formats with differing characteristics are available. The problem therefore lies in selecting a format which is most appropriate not only for a particular use and view of the product, but also for long-term retention.

For this purpose we have developed the Registry/Repository of Representation Information for Engineering (RRoRIfE), a simple preservation planning tool based on RI relating to various characteristics of file formats and their associated conversion software. The premise behind the tool is that an intellectual object can only be faithfully reproduced in a new format or environment if the latter supports properties or characteristics equivalent to those used by the intellectual object in its native format or environment. Furthermore, different tools may be better or worse at re-expressing the constructs of the old format or environment in the constructs of the new.

Underlying RRoRIfE is an ontology of properties, characteristics and constructs of engineering information; since RRoRIfE is at present focused on CAD models, the current ontology includes various two dimensional and three dimensional geometric entities, as well as different compression techniques and forms of metadata. This ontology was derived from a superset of the properties supported by a sample of CAD formats. Two XML schemas have been written using the ontology. The first relates to file formats and describes whether or not the format supports a particular property. An intermediate value of 'partial' support is allowed, to indicate that support is limited in some way; for example, NURBS surfaces may be allowed, but only with 256 or fewer control points. In cases of partial support, explanatory text must be provided.

The second XML schema relates to processing software. For each format conversion the software is able to perform, a record is created as to how well the conversion preserves each property. Four levels of preservation are allowed: 'none' indicating that the property has never knowingly survived the conversion intact (perhaps because the destination format does not support the property); 'good' indicates that the conversion has so far preserved examples of the property sufficiently well that it would be possible to reconstruct the original expression of the property from the new expression; 'poor' is used when tests have found it at least as likely for the property to be corrupted or lost as it is to survive; while 'fair' is used otherwise, alongside an explanatory note.

Figure 5: User Interface to RRoRIfE

Where preservation is less than 'good', it is possible to record whether the property survives in a degraded form, and if so, whether this degradation always happens in a fixed way, a configurable way or an unpredictable way. For example, when moving from a format that supports NURBS to one that only supports tessellating triangles, there may be a fixed algorithm for approximating surfaces, or one may be able to specify how detailed the approximation is.

RRoRIfE reads files in these two XML formats, and uses them to answer simple preservation planning queries. As well as being able to look up the characteristics of individual formats and conversion utilities, RRoRIfE also allows one to select certain characteristics as significant and discover which formats support them. It can generate possible migration pathways between two formats, and given a starting format and set of significant characteristics, it can generate a list of suitable destination formats and conversion pathways. Figure 5 shows the GUI to RRoRIfE.

5. FURTHER WORK

There are several aspects to continuing the work described in this paper, not least that of developing an integral framework for LiMMA and RRoRIfe. One issue currently receiving much attention concerns the persistent identification of geometry between translations from native CAD models into lightweight formats, such that product information can be reliably associated with the same entities in both models. This is also known as the "persistent naming problem" [Marcheix & Pierra, 2002; Mun & Han, 2005], and is not peculiar solely to engineering data. Buneman et al. discuss an analogous problem in the context of curating databases [Buneman et al, 2008].

With respect to RRoRIfe, it is important to accumulate a corpus of RI to enable informed decision making both in terms of the characteristics of conversion software as well as investigating the significant properties of various formats.

Wider consideration of digital curation in PLM includes non-intrusive and automated capture of information as well as issues relating to the selection and appraisal of product lifecycle data.

6. CONCLUSIONS

We have examined the digital curation challenges posed by PLM in engineering and suggested several techniques to improve the robustness of product data to serve the needs of both PLM and long-term accessibility.

Full CAD models tend to have closed, proprietary formats and are difficult to pass around between organisations and the stages in PLM. Lightweight formats provide a more promising approach in that: they have open specifications; they are simpler and have smaller file sizes; they can cater to the need for multiple viewpoints; and restrict access for security purposes. In addition, multilayered annotation allows models to be augmented with much valuable data including that required for preservation. Finally, accumulation of RI facilitates informed decision making with respect to which lightweight formats and conversion software to use both for data exchange within PLM and for accessibility in the longer run.

7. REFERENCES:

Adobe Systems, 2007a. PDF Reference and Related Documentation. Adobe Acrobat SDK version 8.1, http://www.adobe.com/devnet/pdf/pdfs/PDFReference16.pdf (accessed 14th June 2008).

Adobe Systems, 2007b. PRC Format. Version 7094, http://www.adobe.com/devnet/acrobat/pdfs/pdf_reference.pdf (accessed 14th June 2008).

Atlantic Workshop on Long-Term Knowledge Retention (LTKR), 2007. Dept. Mechanical Engineering, University of Bath, UK, 12-13th February 2007.

Ball A., Patel M., McMahon C., Culley S., Green S., 2006. A Grand Challenge: Immortal Information and Through-Life Knowledge Management (KIM), IJDC 2006, Vol. 1, No. 1.

Beagrie N., 2006. Digital Curation for Science, Digital Libraries, and Individuals, IJDC 2006, Vol. 1, No. 1.

Buneman P., Cheney J., Tan W-C., Vansummeren S., 2008. Curated Databases, PODS'08, June 9-12th 2008, Vancouver, BC, Canada

CCSDS 2003. Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-B-1. Blue Book (ISO 14721:2003).

Dassault Systèmes, 2007. 3D XML User's Guide. 0th ed. Version 4.0.

Digital Curation Centre (DCC), 2007. What is Digital Curation? http://www.dcc.ac.uk/about/what/ (accessed 10th June 2008).

DCC Curation Lifecycle Model, 2008. http://www.dcc.ac.uk/docs/publications/DCCLifecycle.pdf (accessed 10th June 2008).

Ding L., Li W.D., McMahon C.A., Patel M., 2006. Lightweight Representations for Product Lifecycle Management, Virtual Concept 26th November – 1st December 2006, Mexico.

ECMA-363, 2007. Universal 3D File Format. 4th Edition.

Extensible Markup Language (XML) 1.1 (Second Edition), 2006. W3C Recommendation. World Wide Web Consortium.

Giaretta, D., 2007. The CASPAR Approach to Digital Preservation, International Journal of Digital Curation, Vol 2 (1) 2007.

IDC white paper, 2008. The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010. Sponsored by EMC, http://www.emc.com/about/destination/digital_universe/ (accessed 14th August 2008).

Initial Graphics Exchange Specification (IGES) v5.3, 1996. ANSI, from US Product Data Association (USPRO), http://www.uspro.org/documents/IGES5-3_forDownload.pdf (accessed 14th August 2008).

ISO 10303 STEP -Standard for the Exchange of Product Model Data, 2005.

Kuny T., 2007. A Digital Dark Ages? Challenges in the Preservation of Electronic Information, 63rd IFLA Council and General Conference, http://www.ifla.org/IV/ifla63/63kuny1.pdf (accessed 14th June 2008).

Kunze, J., 2005. Future-Proofing the Web: What We Can Do Today, 16th September 2005, International Conference on Preservation of Digital Objects, Göttingen, Germany.

Marcheix D. and Pierra G., 2002. A survey of the persistent naming problem, Proceedings of the Seventh ACM Symposium on Solid Modeling and Applications. Saarbrücken, Germany. ACM : New York. pp13–22. ISBN: 978-1-58113-506-0. DOI: 10.1145/566282.566288.

McMahon C., Giess M. and Cully S., 2005. Information management for through life product support: the curation of digital engineering data, IJPLM, Vol. 1, No. 1, pp26-42.

Mun D. and Han S., 2005. Identification of topological entities and naming mapping for parametric CAD model exchanges. International Journal of CAD/CAM Vol. 5, No. 1, 2005.

Open HSF Initiative, 2008. The HOOPS 3D Product Suite, http://www.openhsf.org/docs_hsf/index.html (accessed 14th June 2008).

Patel M. & Ball A., 2007. Challenges and issues relating to the use of Representation Information for the digital curation of Crystallography and Engineering Data, 3rd International Digital Curation Conference, 12-13th December 2007, Washington, D.C.

PREMIS Data Dictionary for Preservation Metadata version 2.0, 2008. Preservation Metadata Maintenance Activity, http://www.loc.gov/standards/premis/ (accessed 14th June 2008).

Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC), Version 1.0, 2007. Center for Research Libraries and RLG Programs.

Significant Properties of Digital Objects, 2008. JISC/BL/DPC Workshop, April 2008

Strodl S., Becker C., Neumayer R., Rauber A., 2007. How to Choose a Digital Preservation Strategy: Evaluating a Preservation Planning Procedure, JCDL'07, 17–22 June, 2007, Vancouver, British Columbia, Canada.

TEI Standoff Markup Working Group, 2003. Stand-off markup. Working paper SOW 06, Text Encoding Initiative.

Thompson, H. S. and McKelvie, D., 1997. Hyperlink semantics for standoff markup of read only documents: The next decade – pushing the envelope, Proceedings of SGML Europe '97. Barcelona. Graphic Communications Association, pp227–229.

University of California, Berkeley, 2000. How much Information, http://www2.sims.berkeley.edu/research/projects/how-much-info/summary.html (accessed 14th August 2008).

UGS, 2005. Open product lifecycle data sharing using XML, http://www.ugs.com/products/open/plmxml/docs/wp_plm_xml_14.pdf (accessed 14th June 2008).

UGS, 2006. JT File Format Reference. Version 8.1, http://www.jtopen.com/docs/JT_File_Format_Reference.pdf (accessed 14th June 2008).

Versprille K., 2005. Dassault Systèmes' strategic initiative: 3D XML for sharing product information. Technology Trends in PLM. Collaborative Product Development Associates.

XGL Working Group, 2006. XGL File Format Specification, http://web.archive.org/web/20060218/http://www.xglspec.org/ (accessed 14th June 2008).

8. ACKNOWLEDGEMENTS

The Digital Curation Centre is funded by the UK's Joint Information Systems Committee (JISC). The KIM Project is supported by the UK Engineering and Physical Sciences Research Council (EPSRC) and the Economic and Social Research Council (ESRC) under Grant Numbers EP/C534220/1 and RES-331-27-0006.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

IMPROVING ARCHITECTURAL DESIGN ANALYSIS USING 3D MODELING AND VISUALIZATION TECHNIQUES

Stefan Boeykens, Mario Santana Quintero, Herman Neuckermans

K.U.Leuven, Department of Architecture, Urbanism and Planning, Kasteelpark Arenberg 1, B-3001 Heverlee (Belgium)
(stefan.boeykens, mario.santana, herman.neuckermans)@asro.kuleuven.be

KEY WORDS: CAD, 3D Modelling, Visualization, Analysis, Reconstruction

ABSTRACT:

In Architectural Design and Education, it is common to refer to prior design cases. This is apparent in a teaching context, where students study existing projects, but also in a professional context, where reflection on existing design examples can inform the designer about possible solutions or as historic reference. With the increased usage of 3D techniques in visualization, simulation and Building Information Modelling, architects nowadays produce more and more designs as 3D models. While these models provide new means to visualize and interrogate the design, much of this potential is left unused, as the models are seldom shared to exchange design information.
This article discusses results from the 3D reconstruction of exemplary building projects and sites from recent history. The reconstructions used widely differing techniques, from regular 3D modelling using CAD and visualization software, to extensive measuring and surveying techniques. These examples illustrate the added value 3D models enable, compared to traditional drawings or photographs. Even the structure and presentation of recent design projects can be improved using diagrams and overlays, capitalizing on the results of the 3D modelling efforts.
In parallel, it is possible to improve and increase information about the design, by adding additional metadata to the 3D model. The “enrichment” of the 3D models make better structured information available, which can in turn, facilitate the retrieval and recovery of such models, when searching or browsing for design information through online repositories.
The combination of these diverse techniques enables an increased accessibility of the inherent design information, which would not be established using each technique as such.

1. INTRODUCTION

1.1 Overview

This article discusses a series of digital reconstructions of architectural projects. The first section summarizes the case studies while the following sections describe conclusions and recommendations on the followed techniques and methodology.

1.2 Context

During the previous years, several exemplary architectural designs have been reconstructed at the Design & Building Methodology research group from the K.U.Leuven Department of Architecture, Urbanism and Planning (Leuven, Belgium). The projects comprise an interesting mixture of historic and modern buildings. They have been elaborated by our research group, but also in the course of master theses projects, where the models served as a medium to perform an analysis of the architectural design and, where appropriate, their historic context.

Different applications and different techniques have been applied, from generic modelling with CAD software to visualization techniques and animation. In most studies, the models were not the final outcome, but they fully served as a basis to create presentations, diagrams and visual analysis drawings.

An important motivation for these case studies lies in their educational value. Students commonly learn about architecture by looking at exemplary architectural design projects. While drawings and photographs are still a very suitable method to present these cases, current modelling and presentation techniques provide increased interaction and embedded information, which benefits the learning experience.

Historic reconstructions allow the recreation of the architectural artefact for different time periods. Modern CAAD techniques present a new arsenal of techniques of modelling, representation and analysis, as described in (Alkhoven, 1991). Models of any chosen time period can be reconstructed and places inside their original context, using a hybrid of resources, from areal photography to historic manuscripts and custom models. Especially in cases where the building is currently demolished or largely renovated or altered, the reconstructed model can be used to provide insight into the evolution of the building or the site.

1.3 Case Studies

The next section presents a summarized overview of conclusions that emerged from a series of case studies, which have been elaborated partly in collaboration with several master students, during the previous years. In most cases, the reconstruction model was used not only for visualization but also to assist an architectural analysis, e.g. on historic reconstruction, circulation patterns or the use of light and daylighting.

The case studies have different characteristics and have utilized different techniques. Information retrieval is often the result of a literature research, but if the project was accessible, on-site measurements have been taken, using a combination of traditional and photogrammetric techniques. In some cases 3D laser scanning techniques have been applied, as described in (Schueremans and Van Genechten, 2007) and (Santana Quintero and Van Genechten, 2007). The visualization outcome is also identified, ranging from 2D CAD drawings to rendered 3D models or even interactive scenes.

The case studies include historic reconstructions, but also 20th century modern buildings.

- Town Hall and Historic City Center (Leuven, Belgium) (Vandevyvere et al, 2007)
- Maison Du Peuple by Victor Horta (Brussels, Belgium)
- Indochina University by Ernest Hebrard (Hanoi, Vietnam)
- Ch. N.D. Du Haut by Le Corbusier (Ronchamp, France)
- Castle (Horst, Belgium)
- Palais Stoclet by Joseph Hoffmann (Brussels, Belgium)
- Vitra Pavillion by Tadao Ando (Wheil-Am-Rhein, Germany)
- National Assembly Hall by Louis I. Kahn (Dhaka, Bangladesh)
- Art Museum by Axel Schultes (Bonn, Germany)
- Castle Boussu by Jacques Du Broeucq (Mons, Belgium)
- Hunting Residence Mary of Hungary, (Mariemont, Belgium) (Vandevyvere et al, 2007)
- Rito Library by Henri Vandevelde (Leuven, Belgium)
- Béguinage Church and Site (Hasselt, Belgium)
- Broodhuys (Brussels, Belgium)
- Church of Saint-James (Leuven, Belgium) (Schueremans & Van Genechten, 2007)
- Palace of Justice by Joseph Poelaert (Brussels, Belgium) (currently in progress)

2. DIFFERENT APPROACHES TO MODEL AND VISUALIZE RECONSTRUCTED BUILDINGS

An important advantage of digital reconstructions and 3D models over regular drawings or photographs is the added information they can represent. Even fairly simple models can be used to create augmented diagrams, overlaying graphics and text on top of rendered images.

2.1 Modelling

Table 1 juxtaposes different modelling techniques. Not a single technique is best fit for both drawings and 3D models with full support for organic or freeform geometry and with included visualization. In all cases, some compromises have to be made, which leads to the need to translate models between different applications.

	Advantages	Disadvantages
Mesh Modelling (e.g. SketchUp)	Quick to model	No 2D drawings No visualization
Generic CAD (e.g. AutoCAD)	2D drawings 3D models Accuracy	Disconnect between drawings and model
Building Information Modelling (e.g. ArchiCAD)	2D/3D developed concurrently Integrated listing and visualization	Limited for freeform geometry External tools for advanced visualization
Digital Content Creation (e.g. 3ds Max)	Freeform models Visualization and animation	No drawings Difficult for accuracy and scale

Table 1: Comparison of Modelling Techniques

Regardless of the chosen application(s), architects, designers or researchers are faced with work flow problems. Data has to be translated between very different systems, often requiring partial remodelling or restructuring of the passed geometry. Transferring more intelligent data, such as parametric assemblies or digital building models is even more problematic. The application of a format such as the Industry Foundation Classes (IFC) is only supported with BIM applications and even then, information gets lost in the translation process. As described in (Mitchell et al, 2007), where the exchange between a design tool and energy analysis was investigated, the IFC model proved to be useful, but still incomplete, while at the same time being hindered by the modelling limitations of the BIM application.

In all cases, the translation process will also be unidirectional, where a model is translated, extended and then, possibly, translated into another tool. There is no way to synchronize these modifications between different applications. For a static reconstruction, this might seem less problematic than the modelling of a design-in-process, but nevertheless, new and updated information might and will become available during a reconstruction project and inserting it into the original model will start the translation process once again.

The choice of a modelling system directly reflects the potential outcome. E.g. the previous table clearly indicates that the need for 2D drawings leaves no choice but to apply CAD or BIM applications. Generic CAD software can use the 2D drawing to generate a 3D model, but consequent changes in the drawing are not reflected to this model. The only method where the 2D drawings can be elaborated alongside the 3D model is the use of BIM software. But the mediocre support for freeform geometry in these applications, make them unsuitable for highly organic architecture. In the reconstruction model of the Maison Du Peuple the ArchiCAD BIM application was applied to create a simplified model of the complete building layout, mainly to derive 2D drawings, while the highly ornamented façade details were recreated in Autodesk VIZ, using parametric lofting techniques. It would have been possible, in theory, to created parametric scripted GDL entities (Nicholson-Cole, 2000) in ArchiCAD as well, but this would have taken considerably more time, especially since there is little repetition in these entities.

A good example of the application of BIM methods can be found in a case study for historic reconstruction of synagogues (Martens et al, 2002). The study suggests that a structured approach can hugely improve the documentation and reconstruction process, which is important with the sometimes delicate nature of these projects.

2.2 Rendering and Visualization

With the potential of defining accurate simulations of materials and lighting, modern visualization applications provide means to create images that are of a similar quality as photographs. The term photo-realistic has been used for quite some time and visualization artists have produced images that can not be distinguished from reality for many years. An excellent example of the quality that can be achieved, even with currently outdated technology, are the reconstructions of unbuilt works of Louis I. Kahn (Kent Larson, 2000).

However, this aspect required extensive experience, which often means that these results were not obtainable by architects or designers, who need visualization as a by-product of the design process. Similarly, creation historic reconstructions up to a level of visual quality that seems to be expected with current technology, often demands the outsourcing of this tasks to artists, leading to results as collected in the books from Ballistic Publishing, such as Exposé (Snoswell, Teo, Eds., 2003) and Elemental (Wade, Snoswell, Eds., 2004).

The integration of improved lighting and material simulation, leading to applications such as Maxwell Render, by Next Limit (http://www.maxwellrender.com), has shown the potential to utilize algorithms which mimic the behavior of light in the real world, rather than using simplified methods to create nice looking images.

In the Vitra Pavillion and Ronchamp Chapel case studies, attempts were made to juxtapose images of the 3D model alongside photographs taken on site. In both cases, the same software was used, in casu Autodesk VIZ, but the availability of more efficient and easier to use rendering engines has displayed great improvements over the last few years, as displayed in Figure 1, which was created in 2003 and Figure 2 which was rendered in 2007.

Figure 1: Vitra Pavillion, picture versus rendering

Figure 2: Picture versus Rendering of Ronchamp Chapel

The comparison mainly proves the achievable quality. The added value, however, is the possibility to create images that are not available with regular means, such as orthographic views, bird-view images or even disassembled sections. Figure 3 shows a rendered orthographic projection of the Indochine University reconstruction.

Figure 3: Orthographic Rendering of Indochine University

Figure 4 displays the dining room in the Palais Stoclet. While this building is still intact, it is not open to the public. The reconstruction was based on available drawings, on-site reference measurements using a Total Station system and images available from literature.

Figure 4: Palais Stoclet interior and open perspective

In this particular example, the added value was not only the different rendered images, but the possibility to derive and deconstruct the model, creating sections or see-through perspectives, which can not be obtained with other methods.

Figure 5 displays a conceptual rendering of the dining room as an open perspective, locating this room in the whole of the building.

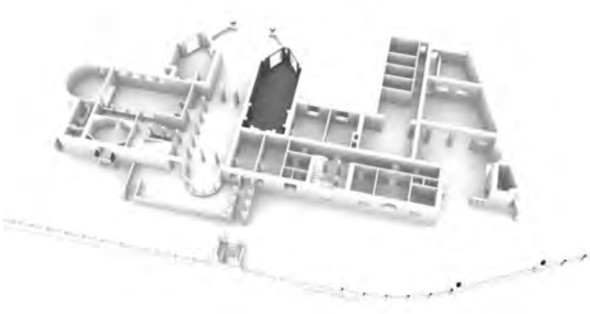


Figure 5: Locating the dining room in Palais Stoclet

2.3 Schemes and diagrams

While photorealism is often a desired outcome of a 3D reconstruction, it is not always required to provide insight. Many of the case studies have used the 3D model to generate not necessarily realistic pictures, but attractive visual representations of the building structure. By overlaying the images with additional annotation, such as arrows, text or colors, it is possible to create visually attractive diagrams, which can provide more insight into the building, with a visual language that might appeal to a wider audience.

Figure 6 displays the Horst Castle twice. The left image displays the different building phases, whereas the right image shows the degree of accuracy from the historic references.

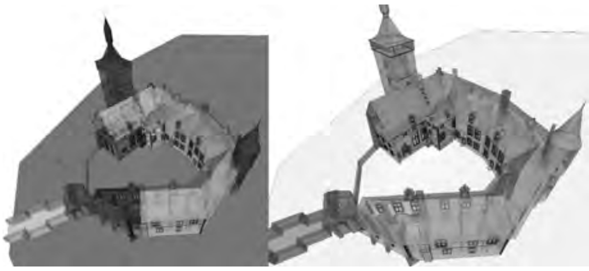


Figure 6: Color-coded model of Horst Castle

Figure 7 shows a similar approach to display the degrees of accuracy as witnessed in the reconstruction of the Hasselt Béguinage Church.

Figure 7: Color-coded model of Hasselt Béguinage Church

The same reconstruction also displayed the church in its environment, for different time periods. Figure 8 compares the 1842 context with the current situation, where only a ruin is left.

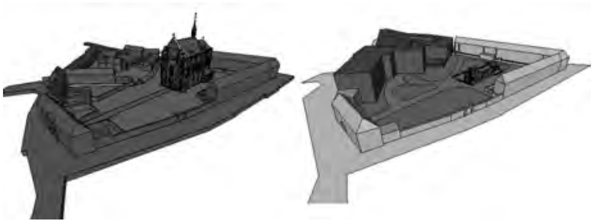


Figure 8: Comparing 1842 and 2007 context

The possibility to augment renderings with additional information is only partially achievable with regular drawings or photographs. Especially in an analysis and educational context, this proves to be very useful.

2.4 Realtime visualization

The Horst Castle and the Hasselt Béguinage models were converted to VRML files (http://www.web3d.org/x3d/vrml), as shown in Figure 9. This allows real-time exploration and gives the possibility to embed interactivity and hyperlinks. While the VRML technology is currently superseded with the X3D initiative, it is still widely supported by many modeling systems and still presents an accessible approach, despite its limitations, such as lack of streaming support.

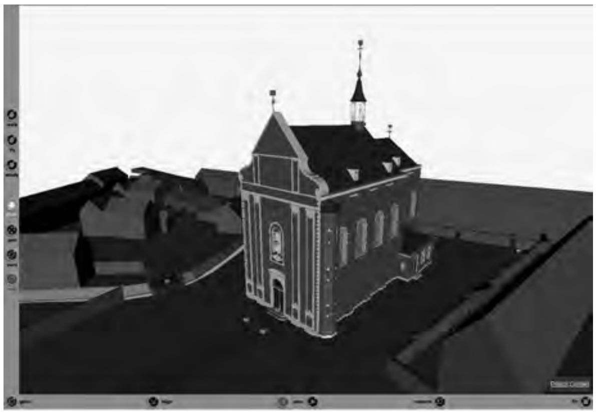


Figure 9: VRML model of Hasselt Béguinage Church

The model of the Art Museum in Bonn was translated to the Unreal game engine (http://www.unrealtechnology.com). 3ds Max was used to transfer the original AutoCAD model into ASE files for import as so-called "static meshes". While the end result was an interactive model, complete with textures and partial shadows, the process was very involved.

The possible end result could ultimately lead to models such as the famous reconstruction of the F. L. Wright Kaufmann House (Falling Water) by 3D artist Kasperg (http://twhl.co.za/vault.php?map=3657) using the Half-Life 2 level editor. This particular model is rich in visual quality, utilizing lights, shadows, textures and even sound to convey an immersive end result. Yet, the process required a considerable amount of effort, bypassing traditional modeling tools and using the fairly primitive level editing software that was developed for the game. This was required to get to a playable and reasonably efficient model, which would not have been possible using common 3D modeling techniques.

In the Stoclet reconstruction, the model was translated into Quest3D (http://quest3d.com) as shown in Figure 10, which was used to generate a self-contained interactive model, allowing the user to walk through the building in realtime, exploring the structure and the different corners of the building.

Figure 10: Quest 3D interactive model of Palais Stoclet

The model was quite elaborate, but still was not a complete reconstruction, leaving several unfinished sections and gaps. They are not immediately noticeable, but can be discovered in the interactive model. There are some important limitations, however. The interactive model requires a Windows system and can only be run on a fairly well equipped machine with a decent graphics adaptor to be able to smoothly display the model.

In the case studies, mainly three different techniques have been studied, with different advantages and disadvantages. They are displayed in Table 2. All systems have decent support for hardware acceleration, using the graphics adapter.

	Advantages	Disadvantages
VRML models	Open Standard Widely supported Interactivity Hyperlinks Cross-platform	No shaders in current viewers No streaming
Game Engines (e.g. Half-Life 2 Source Engine, Unreal Engine)	Cheap application Free level editing Visual Quality (shaders)	Elaborate transfer “Weapons” visible Expensive licensing
Dedicated interactive systems (e.g. Quest 3D, VRContext)	Extensive interactivity Standalone viewer	Expensive Not always cross-platform Complex to use

Table 2: Comparison of realtime techniques

The potential of these systems is often diminished by the large effort it takes to translate an architectural model from a CAD or 3D application into a usable interactive model. The process is unidirectional and many steps take a considerable effort to turn the model into an efficient scene.

Some software firms try to solve this by providing direct exporting modules for a CAD or 3D system, but this usually limits the possibility to insert interactivity and the solutions are often expensive. Examples include Cult3D, TurnTool, EON Reality and VRContext.

3. STRUCTURING THE RECONSTRUCTION

3.1 Methodology

The methodological framework that was applied several times throughout the historic reconstruction case studies is described in (Vandevyvere et al, 2007). It involves surveying, historic investigation, the creation of a “Metafile”, the modelling of the 3D digital model and finally an enabling (multimedia) interface.

3.2 Metafile of resources and accuracies

The Metafile is a tabular listing or a database of all retrieved and referenced resources, from documents, images, manuscripts and books. By summarizing all known historic facts or claimed construction steps, still referencing their source, they can be juxtaposed and compared. While the method of building such a table is quite straightforward, choosing categories or table rows is still investigated on a case-by-case basis. In some cases, this could follow the historic time line, whereas other cases have used different parts of the buildings to categorize the information.

Table 3 displays a small mockup table, indicating the kind of information that can be noted in the Metafile.

build/part	Facts	Sources	Iconography	…
Front House	First stone 1448 … Restoration 1829	V-M constr. Bill … Van Even Descr.	View on Market, 1610 … Survey Plans In situ check 1997	
Roof Structure	Orig. structure 1452-1460 Material oak … Carpenter selected 1452	V-M constr. Bill ... Internal note 1997	In site survey	
…				

Table 3: small fragment of Metafile

Even when the full table or database is created, it is not trivial to format it in a meaningful and readable layout. It would be helpful to create an interactive interface around the table, to be able to create ad-hoc filters and queries, while still being able to place related information side-by-side.

4. AVAILABILITY FOR A WIDER AUDIENCE

Even though the case studies indicated the different techniques that can be applied for modelling and visualization, there is often the need to disseminate and communicate these results to a wider audience, such as students or visitors of an exhibition portal or website.

While it is easy to present rendered images in a text or on a website, it is beneficial to provide some form of interactivity. The case studies have also investigated some possible approaches to enable this.

4.1 Presentation Techniques

The Mariemont reconstruction was used in an art exhibition, in the form of an interactive Flash gallery, on a CD-ROM. A series of rendered images was used to allow visitors to interactively turn around the reconstructed model, without the requirement of a real-time system or loading full 3D models. Most computers have the Flash player installed, which makes the result potentially compatible towards a larger audience.

The Hasselt Béguinage case study presented the result as a website, with hyperlinks to the VRML model. This model in its turn contained several links, embedded inside the interactive scene towards additional references, such as manuscript excerpts, pictures or PDF documents. The model became the interface. This was possible since the model itself had only a low level of detail. The end user has to install a VRML browser plug-in, however, to be able to load the interactive model in a web browser.

While the Bonn Art Museum and the Stoclet Palace have utilized more advanced real-time animation systems, they suffered from a very labor intensive translation effort and high system requirements, making the interactive model only suitable for powerful Windows-based CAD or gaming workstations. The interactivity was also limited to exploring the scene, using gravity and collision detection. The programming of additional interactivity proved to be difficult and would ad an additional level of complexity in the reconstruction. Ironically, the study using the aged VRML format realized better interactivity, using the helper objects in 3ds max, to embed actions such as following an external hyperlink.

The Leuven Town Hall and the Ronchamp Chapel have been translated into STL files and were used to generate physical models, using Stereo Litographic techniques, as shown in Figure 11.

Figure 11: Stereo Lithographic model of Leuven Town Hall

However, the end results took several hours of preparation by an expert, to optimize the model for a faultless output. Moreover, the final model, while intricately beautiful and visible for a non-specialist audience, is very brittle. Touching the model is not feasible in an exhibition context, as the small details can easily be broken.

A potential further exploration is the inclusion of external reference information inside an interactive system. Examples such as Google Earth or Second Life illustrate that there are means to interact between a virtual world and an online community, by connecting content from external sites into the system. With Google Earth, users can create 3D models of sites or buildings and allow users to load them into this world. There are several users investigating means to embed dynamic information into otherwise static environments, e.g. embedding real-time weather or mapping information into Second Life, as presented on the Digital Urban blog (Batty, Smith, 2005).

4.2 Towards retrievable information in an online repository

The creation and the presentation of reconstruction models poses some problems, which are mostly due to the large size of the 3D models, the different applied proprietary file formats and the commonly unstructured models. To make such models usable in content libraries thus presents a series of technical and logistic problems.

The authors are involved in MACE (Metadata for Architectural Contents in Europe). This is a European eContentsPlus project (http://www.mace-project.eu), which investigates the usage of metadata to improve access to architectural content in online repositories (Heylighen et al., 2007). Within this project, an approach was set up to properly classify architectural content, by defining different taxonomies. The architectural domain taxonomy is a combination of common architectural classification systems, describing architectural features, such as function, performance and construction information. The media taxonomy, on the other hand, describes media objects, such as pictures, texts and other digital files. This taxonomy will have to be extended to properly cater for the description of 3D models. Common information about models could be collected, such as the kind of geometry, the amount of polygons, the availability of material and lighting information and so on. Online retrieval of information is enabled by searching through collected metadata, rather than looking at the actual models. This does not directly solve the issues of large file sizes and proprietary formats, but it will facilitate the online recovery of models.

In addition, it is envisaged to make more extensive use of Open Standard formats, such as VRML or IFC, which enables models to become accessible in the future. Through these case studies, the problem of recovering models in proprietary formats is already apparent, even for models which have been created only a few years ago, because of changes in applications and formats. A long-term strategy has to cater for their conversion into open and documented formats, to ensure their availability in the following years. With most design applications evolving into yearly updates, this problem might increase even more in the future. It is important to keep digital reconstructions accessible for the following generations.

CONCLUSIONS

While these case studies have shown the added value of digital reconstructions, for models from any time period, their outcome is still limited by several factors.

The models often become very large and cumbersome to handle. They are also not directly usable for exposition towards a larger audience. And finally, much of the potential of structured information is lost, because the end results were often created in non-architectural applications, such as game engines and visualization applications.

To combine interactive models with embedded architectural information requires the combination of different techniques, such as transferring models to open standards and the addition of metadata to facilitate online retrieval.

Future research could explore these possibilities.

ACKNOWLEDGEMENTS

Figure 1: Leentje Verfaillie
Figure 2: Joris Lobelle
Figure 3: Model: Hoang Phong Lan, Render: Stefan Boeykens
Figure 4, 5, 10: Kristof Lesaffre
Figure 6: Peter Sterckx
Figure 7, 8, 9: Peter Savenay
Figure 11: Model: Han Vandevyvere, Manufacturing: PMA at K.U.Leuven University

REFERENCES

Alkhoven, P., 1991. The Reconstruction of the Past: The Application of New Techniques for Visualization and Research in Architectural History, in: *Computer Aided Architectural Design Futures: Education, Research, Applications* [CAAD Futures '91 Conference Proceedings / ISBN 3-528-08821-4] Zürich (Switzerland), July 1991, pp. 549-566.

Batty, M., and Smith, A.. 2005 Urban Simulacra: From Real to Virtual Cities and Back and Beyond, in: *Architectural Design, Sensing the 21st Century City: The Net City Close-up and Remote*, David Grahame Shane and Brian McGrath (Editors), ISBN 978-0470024188.

Heylighen, A., Neuckermans, H., Wolpers, M., Casaer, M., Duval, E., 2007. Sharing and Enriching Metadata in Architectural Repositories, in: *Predicting the Future* [25th eCAADe Conference Proceedings / ISBN 978-0-9541183-6-5] Frankfurt am Main (Germany), Sep 2007, pp. 401-408.

Larson, K., 2000. *Louis I. Kahn: Unbuilt Masterworks*. The Monacelli Press, Inc., New York, USA, 2000.

Martens, B. and Peter, H., 2002. Developing Systematics Regarding Virtual Reconstruction of Synagogues. In *Thresholds—Design, Research, Education and Practice, in the Space Between the Physical and the Virtual, Proceedings of the 2002 Annual Conference of the Association for Computer Aided Design In Architecture*, pp. 349–356, Pomona (California) USA, Oct 2002. ACADIA.

Mitchell, J., Wong, J. and Plume, J., 2007. Design Collaboration Using IFC, in: *Computer Aided Architectural Design Futures.* Proceedings of the 12th International Conference on Computer Aided Architectural Design Futures (978-1-4020-6527-9) Sydney (Australia), pp. 317-329.

Schueremans, L., Van Genechten, B., 2007. The use of 3D-scanning in structural restoration projects – a case study on the Church of Saint-James illustrating the possibilities and limitations, in: *International Journal Optics and Lasers, special issue related to 3rd Workshop on Optical Measurement Techniques for Structures and Systems*, Optimess 2007, May 2007, Leuven (Belgium), accepted for publication, July 31, 2007.

Santana Quintero, M., Van Genechten, B., 2007. Three-Dimensional risk mapping for anti-disaster recording of historic buildings, in: *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol.XXXVI-5/C53, pp.659-663, ISSN 1682-1750, XXIth International CIPA Symposium, Oct 2007, Athens, (Greece).

Snoswell, M. and Teo, L. (editors) 2003. *Exposé 1 : Finest digital art in the known universe*. Exposé Series. Ballistic Publishing, Australia, first edition.

Vandevyvere H. , Neuckermans H., De Jonge K., 2006. Digital historical reconstruction; Case studies of an interdisciplinary task in: *International Journal of Design Sciences & Technology*, 13(1), pp. 51-65.

Wade, D. and Snoswell, M. (editors). 2004 *Elemental : The World's best Discreet art*. Elemental Series. Ballistic Publishing, Australia, first edition.

Nicholson-Cole, D. 2000. *The GDL Cookbook*. Marmelade Graphics, Nottingham (UK), 3d edition.

Digital reconstruction and 3D Modeling

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

3D LINE DRAWING FROM POINT CLOUDS USING CHROMATIC STEREO AND SHADING

Á. Rodríguez Miranda[a], J. M. Valle Melón[a], J. M. Martínez Montiel[b]

[a] Laboratorio de Documentación Geométrica del Patrimonio, Grupo de Investigación en Arqueología de la Arquitectura, Universidad del País Vasco-Euskal Herriko Unibertsitatea (Spain), http://www.ldgp.es
(alvaro_rodriguez, jm.valle)@ehu.es
[b] Departamento de Informática e Ingeniería de Sistemas, Universidad de Zaragoza (Spain). josemari@unizar.es

KEY WORDS: point cloud, 3D line drawing, Chromadepth, chromatic stereo, heritage

ABSTRACT:

Terrestrial laser scanners (TLS) have become a common tool in geometric documentation of heritage. As output, they can produce point clouds along with thematic information, such as reflectance or photographic texture. Some of their drawbacks are that the products which have been directly obtained from raw point clouds imply massive data handling. Besides, standard products demanded by organizations and technicians, such as wireframe models and line drawings, are not readily computed from them; therefore, human supervised point cloud processing is still mandatory.
Our contribution enriches the point cloud thematic component using shading and Chromadepth® information, which allows to enhance the point cloud surface details and to improve the identification and drawing of the lines defining the elements over the point cloud.
In order to illustrate the utility of these thematic layers, we describe, step by step, a simple procedure to draw 3D lines from point clouds.

1. INTRODUCTION

POINT CLOUDS, enriched with thematic information such as laser reflectance or photographic texture, have found widespread usage in several engineering fields; among them, the geometric documentation of heritage can be highlighted. The ability to readily acquire models and products, such as orthoimages or sections, has contributed to its acceptance. Nowadays, however, some of the commonly demanded, such as line drawings, are not readily available from point clouds.

It is interesting to emphasize that, besides the drawing, the extraction of edges from the point cloud within the three-dimensional space has been applied in many areas, such as: image segmentation, 3D scan alignment or photographic image alignment with respect to the point cloud. We are not reporting those works exhaustively, but focusing on those which are relevant for our purpose of line drawing over point clouds.

The high quality line drawing demanded in cultural heritage makes the human interaction with the 3D model unavoidable. As the interaction interface is a computer screen, the operator does not interact directly with the 3D point cloud but with its 2D projections along with the available thematic information on the computer screen. Figure 1 details the interaction process, 3D information is projected in 2D enriched with thematic information to produce the image on the screen, the operator makes the drawing on the computer screen, and hence the line drawing is backprojected on the 3D point cloud, yielding finally the 3D line drawing model.

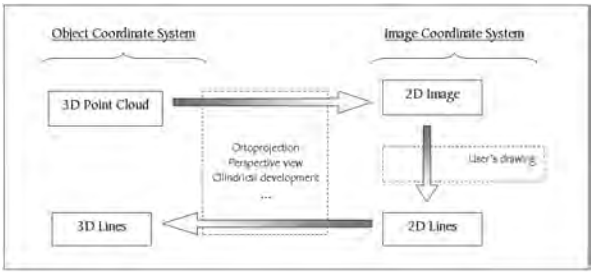


Figure 1: Manual line drawing workflow

Line drawing is highly dependent on the thematic information available. Sometimes this information is degraded during the alignment of several scans, specially when the coded information is laser reflectivity that varies according to the distance and orientation with respect to the imaged object. Because of that we propose to synthesize thematic information from the 3D shape of the object. In our experience we have found particularly useful shading and Chromadepth information.

In the case of the shading, the user can select the relative location of the point cloud with respect to the light source, and hence making visible details of the 3D point cloud worth to direct the drawing.

The use of the Chromadepth® system (this technology is owned by American Paper Optics, Inc.) allows to visualize depth from a single 2D image under a controlled point of view. This visualization of the 3D information has also proved quite effective in guiding the line drawing process in order to achieve the high quality results demanded in heritage documentation applications.

2. INTERFACING 3D POINT CLOUDS BY 2D PROJECTIONS

In general the data format for a 3D raw point cloud is (X,Y,Z,N_D), where the first three values are the coordinates defining the 3D location of a point, and N_D is the thematic vector containing visualization information for the 3D point. Fields typically defining the thematic components are:

- Data captured with the sensor directly, like reflectivity or distance.
- Assigned data obtained from external sources: such as the radiometric information gathered with a digital camera normally located close to the 3D sensor central point. There are even (Reulke, 2006) some sensors like the PMD (Photonic Mixer Device) cameras which can record image and range at the same time.
- Calculated data, for instance: curvature, shading or Chromadepth.

In (Ardissone et al., 2005) a multilayer approach is illustrated by showing the reflectivity image (obtained from the scanner), the range image (calculated, while the viewpoint has been shifted) and the photographic image (from a different source) all together.

It has to be noted that, in general, the point cloud is composed from a set of 3D scans gathered from rather different sensor locations, and it has to be possible to compute the two-dimensional projection from any point of view, by any kind of projection and showing any thematic data. In general, we can apply the perspective projection to compute the 2D image from the point cloud. Nevertheless, for the considered points clouds, points correspond to the surface of the object and hence, it could be possible to represent —at least locally— the point cloud by means of a 2D parameterisation, which implies that the projection method can be extended to include perspectives, orthographic views or cylindrical developments. The concept has been extended to fitted surfaces like the series of cylinders and planes as proposed in (Bonino et al., 2005). This extended range of projections allows selecting the one where the thematic information is shown clearly for the line drawing

According to (Ressl et al., 2006), the interest of merging scanner data and photographic image was seen right from the beginning. Given that the image resolution is usually greater than the point density, one possible solution consists in orientating the image and adding a supplementary channel with the distance in each pixel from the origin of the photograph. The user works over the two-dimensional image but the output is in the three-dimensional object space (Alshawabkeh and Haala, 2004; Bornaz and Dequal, 2004; Abdelhafiz et al., 2005; Neubauer et al., 2005).

Nonetheless, the solution we propose, regarding the photographic texture, consists in assigning the thematic values to the points; as it was said above, it is common for the photographic image to have greater resolution than the point cloud, therefore, some information is lost in this process. This drawback can be eluded, to some extent, by adding new 3D points obtained by interpolation in the photographic image and, as points store attributes, these new points can be ticked so that they are used only if the generated two-dimensional image, over which the user works, is compatible with the photographic one.

3. LINE DRAWING FROM POINT CLOUDS

All the algorithms from image analysis might be used on the 2D thematic images we are working with, moreover, as it is possible to emulate user's criteria when extracting edges, it will be feasible to define automatic algorithms. The use of these techniques in photogrammetry was set by (Schenk, 2000). However, it is still an open issue that does not produce the required accuracy for heritage documentation.

A review of the related works where linear elements are employed illustrates that they fit the scheme previously set (figure 1): (Dorminger and Briese, 2005) represent the shading image on a cylindrical development, (Briese, 2006) shows more examples including reflectivity images, in (Deveau et al., 2005a) the profiles of some surfaces of revolution are extracted from the range image, orientation and curvature images can be found in (Martínez et al., 2005) and, even if it has been designed for meshes, the automatic contour drawing method from normal vectors of (Boehler et al., 2003) can be tailored for point clouds without meshing as well, since there are algorithms to calculate the normal vectors directly from the point as it is described below.

Obviously, the presented scheme can be extended (figure 2) by obtaining several two-dimensional images from different viewpoints (what allows the conventional stereoscopic vision) or the simultaneous management of several thematic layers by means of multi-criterion algorithms for the automatic edge extraction. In (Alshawabkeh et al., 2006) the linear elements from images generated by the scanner (range and reflectivity) are compared with the elements extracted from the photographic images to improve the geometric relationship between photography and scanner data. On the contrary, the outputs from different layers are fused in a unique final edge map, (Deveau et al., 2005b).

For automated processing, the need of a view no longer exists, so the scheme (figure 2) has also the possibility of obtaining the lines directly from the three-dimensional space. In (Gross and Thoennenssen, 2006), a calculated 3D neighbourhood for each point is used to obtain the eigenvalues which permit to identify the edge points. Then, in a second step, these points are converted into line elements. Another example can be found in (Lerma and Biosca, 2004); although the last conversion to lines is not performed, they recognize the edge points by analysing the curvature directly calculated in the 3D space.

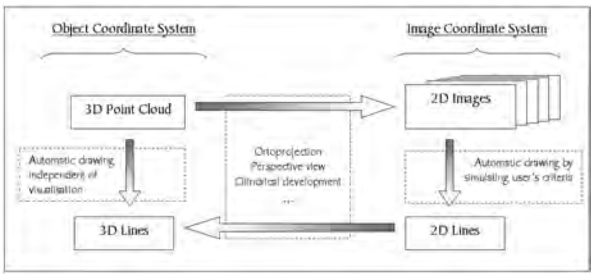


Figure 2: Extented workflow including the automation process

4. LINE DRAWING OVER A ORTHOGRAPHIC PROJECTION. AN EXAMPLE OF IMPLEMENTATION

This section is devoted to the development of a detailed example of line drawing for the orthographic projection. More than a conceptual approach, it is intended as an explanatory example describing step by step how to obtain the three-dimensional lines.

4.1 From 3D point cloud to 2D image

Point clouds are lists of coordinates (X,Y,Z,N_D) where N_D means the thematic component, gray scale, reflectivity or colour, to cite a few.

According to the general outline presented in the introduction, a relationship between the tridimensional object coordinates and the coordinates (column, row) must be defined:

$$\begin{cases} column = f(X,Y,Z) \\ row = g(X,Y,Z) \end{cases} \tag{1}$$

Where *f(X,Y,Z)* and *g(X,Y,Z)* mean whatever mathematic function. The lines will be drawn over these bidimensional coordinates. Then, an inverse relationship will provide the tridimensional coordinates of each pixel. These corresponding functions are indicated by *h(column,row)*, *j(column,row)* and *k(column,row)*:

$$\begin{cases} X = h(column,row) \\ Y = j(column,row) \\ Z = k(column,row) \end{cases} \tag{2}$$

As a practical application, we will develop the obtaining of an orthographic view, similarly to (Jansa et al., 2004), defined by the vector **u**(X_0,Y_0,Z_0) from which, the angles H (azimuth from the Y axis in direct direction, that is, anti-clockwise) and V (colatitude, angle from the Z axis) are calculated as:

$$\tan H = \frac{X_0}{Y_0} \;;\; \tan V = \frac{\sqrt{X_0^2 + Y_0^2}}{Z_0} \tag{3}$$

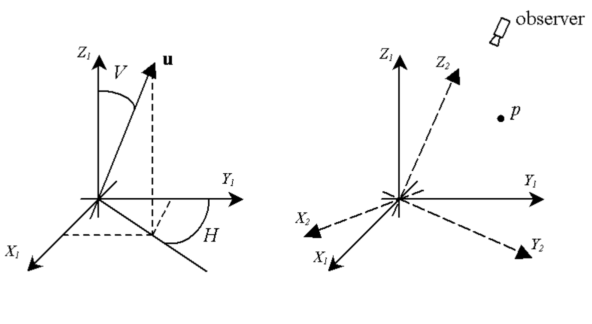


Figure 3: On the left, the vector **u** is drawn in the coordinate system 1; on the right, the new coordinate system 2, whose Z axis follows the **u** vector

A new coordinate system is defined. It has the same origin but its Z axis follows the same direction as vector **u**. These angles permit converting the object coordinate system (X_1,Y_1,Z_1) to the system (X_2,Y_2,Z_2):

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos V & \sin V \\ 0 & -\sin V & \cos V \end{bmatrix} \cdot \begin{bmatrix} \cos H & \sin H & 0 \\ -\sin H & \cos H & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \tag{4}$$

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} \cos H & \sin H & 0 \\ -\cos V \sin H & \cos V \cos H & \sin V \\ \sin V \sin H & -\sin V \cos H & \cos V \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \tag{5}$$

In this new sytem, the observer is placed at infinity in the Z_2 axis so, for each point (p), the higher the value in Z_2, the closer to the observer. For the searched purpose, a third rotation orienting the XY axes towards any specific direction is not needed.

In order to obtain the image, integer numbers for coordinates are needed (X_2,Y_2). Hence, they are rounded to a defined size (1 cm, 5 mm, …). This size is a key value; therefore, it must be chosen carefully to obtain the better resolution without leaving gaps between pixels with data. Once the resolution –*r*- is defined, the new set of coordinates is defined:

$$\begin{cases} X_3 = round\left(\dfrac{X_2}{r}\right) \\ Y_3 = round\left(\dfrac{Y_2}{r}\right) \\ Z_3 = Z_2 \end{cases} \tag{6}$$

We go through the list of points (X_1,Y_1,Z_1) and calculate their (X_3,Y_3,Z_3) coordinates. At the same time, we generate two two-dimensional arrays: $\mathbf{M}(X_3,Y_3)=Z_3$ which stores the highest values (the closest to the observer) of Z_3 for each pair (X_3,Y_3), that is to say, the Z-buffer, and $\mathbf{N}(X_3,Y_3)=N_D$ which stores the digital level held by the point with a bigger value of Z_3. When this is done, we obtain a reduced version of the point cloud with a single point for each pixel, the one with a higher value of Z_3 among all those that have the same (X_3,Y_3) coordinates. It is also necessary to store the extreme values of every coordinate ($X_{3min},X_{3max},Y_{3min},Y_{3max},Z_{3min},Z_{3max}$).

We propose to use this rather simple method to select the image point when several 3D point are viewed in the same pixel. Some examples for more elobarated methods can be found in (Azariadis and Sapidis, 2005), (Liu et al., 2006), (Ressl et al., 2006), in any case, the important point is to establish the existence of a single value in each pixel.

Some special features of the computer language must be taken into account: the source of the image can be either the pixel (0,0) or (1,1), and it is usually placed at the upper-left corner and the Y values (rows) grow downwards; on the contrary, the origin of the image in the (X_3,Y_3) system is the lower-left corner, and the Y values increase upwards from (0,0) source pixel.

$$\begin{cases} column = X_3 - X_{3\min} \\ row = Y_{3\max} - Y_3 \end{cases} \tag{7}$$

The corresponding colour in each pixel comes from the value N_D. Besides the image file, a text file stores the values of resolution, the bounds obtained when scanning the point cloud, the image size (number of rows n_{row} and columns n_{col}) and the values from the array **M**.

Figure 4 presents an axonometric view of a simplified point cloud and the orthoimage calculated according to the described method (the grey scale shows reflectivity values).

Figure 4: On the left, 3D view of the point cloud (decimated); on the right, generated orthoimage.

4.2 Line drawing in 2D and backprojection 2D to 3D point cloud

Then, we can draw the two-dimensional lines of the interesting elements over the image after having inserted it in a CAD software (figure 5, left).

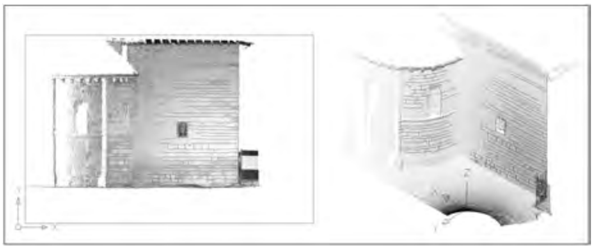


Figure 5: 2D lines over the orthoimage (left), and the same lines after proyecting them over the point cloud (right)

Lines are series of coordinates (x_{CAD},y_{CAD}) with their topology, that is, the links between them. Once we have the two-dimensional lines, they are transformed into the object space by obtaining the corresponding three-dimensional coordinates (X_1,Y_1,Z_1) of each (x_{CAD},y_{CAD}) pair and by keeping their topology.

If the image file has been inserted arbitrarily (place and scale) in the CAD space, we will need to know the coordinates of two corners to transform from (x_{CAD},y_{CAD}) to (*column,row*). For instance, the lower-left (x_{LLCAD},y_{LLCAD}) and the upper-right (x_{URCAD},y_{URCAD}) one. Values must be integer.

$$\begin{cases} column = \text{int}\left[\dfrac{n_{col}\left(x_{CAD} - x_{LLCAD}\right)}{x_{URCAD} - x_{LLCAD}}\right] \\ row = \text{int}\left[\dfrac{n_{row}\left(y_{URCAD} - y_{CAD}\right)}{y_{URCAD} - y_{LLCAD}}\right] \end{cases} \quad (8)$$

In this equation, the origin has been set to (0,0), so the transformations follow like that:

$$\begin{cases} X_3 = X_{3\min} + column \\ Y_3 = Y_{3\max} - row \end{cases} \quad (9)$$

From the (X_3,Y_3) values, we can find in the array **M** what is the corresponding Z_3 value. Then, the coordinates (X_2,Y_2,Z_2) are computed:

$$\begin{cases} X_2 = X_3 \cdot r \\ Y_2 = Y_3 \cdot r \\ Z_2 = Z_3 \end{cases} \quad (10)$$

Finally, the three-dimensional coordinates in the object space are obtained by transposing the array used in (5), so we can go from (X_2,Y_2,Z_2) to (X_1,Y_1,Z_1) as follows:

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = \begin{bmatrix} \cos H & -\cos V \sin H & \sin V \sin H \\ \sin H & \cos V \cos H & -\sin V \cos H \\ 0 & \sin V & \cos V \end{bmatrix} \cdot \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} \quad (11)$$

Figure 5 shows the 2D lines drawn over the image in the CAD system and the same lines in the object three-dimensional space over the point cloud.

This process can be repeated with the same point cloud from different directions (vector **u**) until it completely covers the elements to draw. The generated images are orthoimages so they can be useful for cartographic purposes as they are metric outputs.

5. SYNTHESIZING SHADING THEMATIC LAYER

If a reliable N_D is not available, line drawing in 2D would be rather difficult and inaccurate. This simuation migth seem unlikely as most of the sensors capture both intensity and range data and the former can be displayed like a grey-scale image, but this information is not readily available after having merged several scans gathered from different viewpoints. The reason is that N_D value depends on the distance, the incidence angle and the observed material; as a result, if the final image is obtained just by averaging values, most of the texture will be blurred. The photographic texture might offer richer information; however, external lighting conditions, shadows and radiometric differences might degrade this information.

In poor N_D cases, we can resort to compute an artificial shading to show the relief of the objects represented by the point cloud. Considering the classic *lambertian reflection model* (for instance, in (Folley et al., 1990)) in which the brightness value in each point depends only on the cosine of the angle (θ) between two vectors: the normal vector (**n**) to the surface from the point itself and the vector which defines the direction of the lighting source (**l**). One important feature of this model is that the value is independent of the position of the user, hence, it can be assigned to the point (figure 6, left).

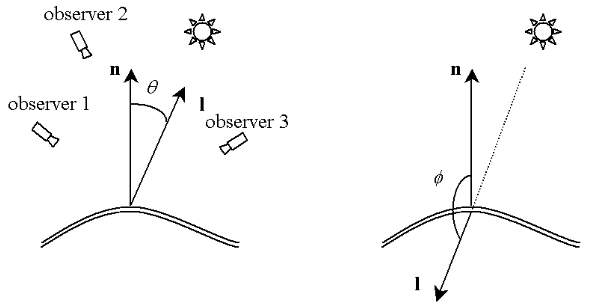


Figure 6: In the Lambertian reflection model (on the left), brightness depends only on the angle (θ), no matter where the observer is. The proposed model (on the right) follows the Lambertian model but considering θ values ranging in the [0, π] interval

The model we are to use is a slight variation (figure 6 right). First, the lighting vector is defined form the source to the surface. In the lambertian model, the points whose angles are greater than π/2 are not illuminated, whereas we will permit all cases in order to assing values to all the point cloud.

If we have two vectors **v**(X_v,Y_v,Z_v) and **w**(X_w,Y_w,Z_w), the cosine of the angle between them is calculated by the following expression:

$$\cos\phi = \frac{X_v X_w + Y_v Y_w + Z_v Z_w}{\sqrt{X_v^2 + Y_v^2 + Z_v^2}\sqrt{X_w^2 + Y_w^2 + Z_w^2}} \quad (12)$$

The user selects the lighting source location but the normal to the point cloud surface has to be computed. There are many methods to calculate the normal vector. In general, they are either based on the definition of local surfaces, or they mesh the point cloud. Anyhow, at each point, normal depends on neighbour points. The stages of this process are (Ou Yang and Feng, 2005):

a) Firstly, for each point, it is decided what points in its neighbourhood are to be used.
b) Secondly, these points are employed to estimate the normal vector.
c) Finally, the inner and outer directions have to be established.

Setting which points are the nearby ones in the object space, in the coordinate system (X_1,Y_1,Z_1), can be difficult; however, in the system (X_3,Y_3,Z_3) it is immediate since the points are regularly arranged along *X* and *Y* coordinates.
We use a method similar to the one proposed by (Yokoyama and Chikatsu, 2006) to calculate the normal vector. For each pixel, a mean plane is calculated using their nearby pixels (with a neighbourhood of 1 pixel, 2, 3, …). The equation of a plane including the point $p(X_P,Y_P,Z_P)$ is:

$$A(X - X_P) + B(Y - Y_P) + C(Z - Z_P) = 0 \quad (13)$$

Once the plane has been defined, the components of the normal line (*a*,*b*,*c*) should meet the following condition:

$$\frac{a}{A} = \frac{b}{B} = \frac{c}{C} \quad (14)$$

When setting the normal from a plane, the line is obtained but it is also necessary to define the inner and outer directions, to solve this, we set *C*=1; this way, the vector is oriented towards the observer.

A loop goes through all the elements in array **M**. For each one, the coordinates (*X*,*Y*,*Z*) are obtained, and there is a new loop that goes through the nearby points. With each neighbour point we calculate the difference of coordinates and with the summations we pose the following equation which permits obtaining the parameters *A* and *B* of the plane:

$$\begin{bmatrix} \sum \Delta X^2 & \sum \Delta X \Delta Y \\ \sum \Delta X \Delta Y & \sum \Delta Y^2 \end{bmatrix} \cdot \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \sum \Delta X \Delta Z \\ \sum \Delta Y \Delta Z \end{bmatrix} \quad (15)$$

After having the values *A*, *B* and *C*, we calculate the components of the normal unit vector:

$$\begin{cases} a = \dfrac{A}{\sqrt{A^2 + B^2 + 1}} = X_{N3} \\ b = \dfrac{B}{\sqrt{A^2 + B^2 + 1}} = Y_{N3} \\ c = \dfrac{1}{\sqrt{A^2 + B^2 + 1}} = Z_{N3} \end{cases} \quad (16)$$

In these identities the coordinates (*X*,*Y*,*Z*) are marked with the sub index -N- to indicate that it is the normal vector and with a -3- because they are in the coordinate system (X_3,Y_3,Z_3).

On the other hand, the lighting vector (subindex -L-) will be defined in the object coordinate system, that is (X_1,Y_1,Z_1), to calculate the angle, both of them have to be in the same system. Consequently, it is necessary to transform the lighting vector to the system (X_3,Y_3,Z_3):

$$\begin{bmatrix} X_{L3} \\ Y_{L3} \\ Z_{L3} \end{bmatrix} = \begin{bmatrix} X_{L2} \\ Y_{L2} \\ Z_{L2} \end{bmatrix} = \begin{bmatrix} \cos H & \sin H & 0 \\ -\cos V \sin H & \cos V \cos H & \sin V \\ \sin V \sin H & -\sin V \cos H & \cos V \end{bmatrix} \cdot \begin{bmatrix} X_{L1} \\ Y_{L1} \\ Z_{L1} \end{bmatrix} \quad (17)$$

If both of them are unit vectors (if not they must be normalized), the angle can be calculated by the equation (12):

$$\cos\phi = X_{N3}X_{L3} + Y_{N3}Y_{L3} + Z_{N3}Z_{L3} \quad (18)$$

The normal vector has its origin in the point to analyse but the lighting vector comes from a external point, so the greatest value happens when $\cos\phi = -1$ (angle π, vectors with opposite directions), and the lowest when $\cos\phi = 1$ (angle 0, the same direction); fitted to the range 0-255 we will supply the array **N**(X_3,Y_3)=N_D with the digital level.

$$N_D = 255\frac{1 - \cos\phi}{2} \quad (19)$$

Fig. 7. Example of an image obtained directly from the reflectivity values after having merged several scans (left) and the shading image (right)

6. SYNTHESIZING CHROMATIC THEMATIC LAYER

Another method to improve the usage of N_D, is the Chromadepth® system. The advantage of this system is that it does not use two separated images, what would require two coordinated pointers to identify a single point in the three-dimensional space. In Chromadepth®, the sensation of relief is obtained by modifying the colour; therefore, it is not suitable for real colour images because the original colour will disappear, but it is very interesting for grey scale images.

This system uses glasses with special prisms which diffract with a different angle according to the wavelength, so hot colours (red) are showed closer than the cold ones (blue). This stereo is easy to implement in the described process because we have the whole range of distances stored in the array **M**.

There are some chromatic scales which work well in Chromadepth®. A very simple one divides the range in four flights sorted by the distance to the observer from the closest to the furthest. The next table shows how each chromatic channel changes (values from 0 to 1), (Schumann and Schoenwaelder, 2002).

	Red –R-	Green –G-	Blue –B-
flight 1	1	from 0 to 1	0
flight 2	from 1 to 0	1	0
flight 3	0	1	from 0 to 1
flight 4	0	from 1 to 0	1

Table 1. Multipliers of the chromatic channels in the Chromadepth® image

The Chromadepth® image makes the sensation of relief, but the objects are not sharply defined. The best option is merging it with the shading image. In figure 8 the chromatic image, the shading and the merged ones can be observed.

Figure 8: The chromatic image (centre) is the sum of the Cromadepth® image (left) and the shading image (right)

To obtain the merged image, we have both arrays $\mathbf{M}(X_3,Y_3)=Z_3$ and $\mathbf{N}(X_3,Y_3)=N_D$. Each pair of coordinates (X_3,Y_3) provides a value of Z_3; from this value we obtain a chromatic combination RGB. On its part, the array **N** supplies the digital level N_D. The new chromatic values R'G'B' can be obtained according to:

$$\begin{cases} R' = N_D \cdot R \\ G' = N_D \cdot G \\ B' = N_D \cdot B \end{cases} \tag{20}$$

Figure 9 shows four orthoimages of a sculpture over which a set of lines (in black) has been drawn. In figure 10 these lines have been projected over the three-dimensional model.

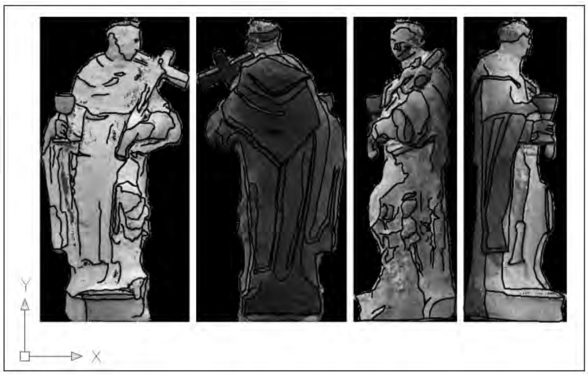


Figure 9: 2D drawing over the Cromadepth® image

Figure 10: 3D lines over the point cloud

It must be emphasized that this kind of stereo permits seeing relief, so the ability to extract information is enhanced; nevertheless, the stereoscopy is not homogeneous through the whole range and the colour range seems strange to the sight. With respect to the comfort and the ability to detect some elements, this method is not comparable with the polarized glasses used in photogrammetry, it should be considered more like a variation of the well-known red-blue anaglyphs.

7. CONCLUSIONS

Line drawing is one of the key products used in heritage documentation. The production of a line drawing from a 3D point under the current automated methods does not meet the quality standards required in heritage documentation, so human interaction is mandatory. We have described the interaction between the user and the 3D point cloud by means of 2D computer display providing an example of detailed equations both for the projection from 3D to 2D, and the back projection from 2D to 3D. We believe this compact coding of the process helps in understanding this interaction with 3D point clouds in heritage documentation production.

One of the key factors in producing 3D line drawings from 2D projections is the availability of a thematic component, normally containing the photometric information guiding the 2D line drawing step. When the thematic component available is poor, it can be enriched by producing synthetic thematic information coming from 3D information. We have focused on shading and in Chromadepth®. Practical line drawing experience has proved this additional thematic information rather useful for producing 3D line drawing from 3D points with poor thematic information.

Based on the proposed model for 3D point cloud interaction, our future work is being directed to automate the line drawing in 2D in order to improve productivity in cultural heritage documentation production.

8. REFERENCES

Abdelhafiz A., Riedel B., Niemeier W., 2005. Towards a 3D true colored space by the fusion of laser scanner point cloud and digital photos. *Proceedings of the ISPRS Working Group V/4 Workshop "3D-Arch 2005, Virtual Reconstruction and Visualization of Complex Architectures" Mestre-Venice, Italy. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, volume XXXVI, part 5/W/17.*. 7 pages. http://www.commission5.isprs.org/3darch05/pdf/31.pdf [Accessed: 28th May 2008]

Alshawabkeh Y., Haala N., 2004. Integration of digital photogrammetry and laser scanning for heritage documentation. *XXth ISPRS Congress, Istambul, Turkey. Volume 35, part B.* pages: 424-429. http://www.isprs.org/istanbul2004/comm5/papers/590.pdf [Accessed: 28th May 2008]

Alshawabkeh Y., Haala N., Fritsch D., 2006. 2D-3D Feature extraction and registration of real world scenes. *ISPRS Commission V Symposium "Image Engineering and Vision Metrology", Dresden, Germany. Volume XXXVI, part 5.* 6 pages. http://www.isprs.org/commission5/proceedings06/paper/ALSH_628.pdf [Accessed: 28th May 2008]

American paper optics inc., 2003. Chromadepth. http://www.chromatek.com [Accesed: 28th May 2008]

Ardissone P., Bornaz L., Lo Turco M., Vitali M., 2005. The relief of the Porta Palatina: a comparison between different survey methodologies and representations. *CIPA XX International Symposium, Torino, Italy*. 5 pages. http://cipa.icomos.org/fileadmin/papers/Torino2005/86.pdf [Accesed: 28th May 2008]

Azariadis P.N., 2004. Parameterization of clouds of unorganized points using dynamic base surfaces. *Computer-Aided Design 36*: 607-623.

Azariadis P.N., Sapidis N.S., 2005. Drawing curves onto a cloud of points for point-based modelling. *Computer-Aided Design 37*: 109-122.

Boehler W., Boehm K., Heinz G., Justus A., Schwarz Ch., Siebold M., 2003. Documentation of stone age artifacts. *CIPA XIX International Symposium, Antalya, Turkey*. 5 pages. http://cipa.icomos.org/fileadmin/papers/antalya/135.pdf [Accesed: 28th May 2008]

Bonino R., Gasco G., Massa A., Rossi S., 2005. Photogrammetric methods applied to the representation of curve facades: an experiment conduced onto Palazzo Carignano in Turin. *CIPA XX International Symposium, Torino, Italy*. 6 pages. http://cipa.icomos.org/fileadmin/papers/Torino2005/130.pdf [Accesed: 28th May 2008]

Bornaz L., Dequal S., 2004. The solid image: an easy and complete way to describe 3D objects. *XXth ISPRS Congress, Istambul, Turkey.* pages: 183-188. http://www.isprs.org/istanbul2004/comm5/papers/546.pdf [Accesed: 28th May 2008]

Briese C., 2006. Structure line modelling based on terrestrial laserscanner data. *ISPRS Commission V Symposium "Image Engineering and Vision Metrology", Dresden , Germany. Volume XXXVI, part 5*. 6 pages. http://www.isprs.org/commission5/proceedings06/paper/1266_Dresden06.pdf [Accesed: 28th May 2008]

Deveau M., Letellier G., Paparoditis N., 2005a. Automating the extraction of revolution objects from single laser scans of architectural scenes. *CIPA XX International Symposium, Torino, Italy*. 4 pages. http://cipa.icomos.org/fileadmin/papers/Torino2005/746.pdf [Accesed: 28th May 2008]

Deveau M., Paparoditis N., Pierrot-Deseilligny M., Chen X., Thibault G., 2005b. Strategy for the extraction of 3D architectural objects from laser and image data acquired from the same viewpoint. *Proceedings of the ISPRS Working Group V/4 Workshop "3D-Arch 2005, Virtual Reconstruction and Visualization of Complex Architectures" Mestre-Venice, Italy. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, volume XXXVI, part 5/W/17.*.5 pages. http://www.commission5.isprs.org/3darch05/pdf/18.pdf [Accesed: 28th May 2008]

Dorminger P., Briese C., 2005. Advanced geometric modelling of historical rooms. *CIPA XX International Symposium, Torino, Italy*. 6 pages. http://cipa.icomos.org/fileadmin/papers/Torino2005/759.pdf [Accesed: 28th May 2008]

Foley J.D., Van Dam A., Feiner S.K., Hughes J.F., 1990. *Computer Graphics, principes and practice* (2nd edition). Section 16.1 *Illumination Models*. pages: 722-726. Adison-Wesley Publising Company. Reading, Massachussets. 1175 pages.

Gross H., Thoennessen U., 2006. Extraction of lines from laser point clouds. *ISPRS Commission III Symposium "PCV'06 Photogrammetric Computer Vision", Bonn, Germany*. pages: 86-91. http://www.isprs.org/commission3/proceedings06/complete_proceedings/final-proceeding.pdf [Accesed: 28th May 2008]

Jansa J., Studnicka N., Forkert G., Haring A., Kager H., 2004. Terrestrial laserscanning and photogrammetry – acquisition techniques complementing one another. *XXth ISPRS Congress, Istambul, Turkey.* 6 pages. http://www.isprs.org/istanbul2004/comm5/papers/184.pdf [Accesed: 28th May 2008]

Lerma J.L., Biosca J.M., 2005. Segmentation and filtering of laser scanner data for cultural heritage. *CIPA XX International Symposium, Torino, Italy*. 6 pages. http://cipa.icomos.org/fileadmin/papers/Torino2005/896.pdf [Accesed: 28th May 2008]

Liu Y-S., Paul J.C., Yong J-H., Yu Pi-Q, Zhang H., Sun J-G, Ramani K., 2006. Automatic least-squares projection of points onto point clouds with applications in reverse engineering. *Computer-Aided Design 38*: 1251-1263.

Martínez J., Finat J., Fuentes L.M., Gonzalo M., Viloria A., 2005. A coarse-to-fine curved approach to 3D surveying of ornamental aspects and sculptures in façades. *CIPA XX International Symposium, Torino, Italy*. 6 pages. http://cipa.icomos.org/fileadmin/papers/Torino2005/441.pdf [Accesed: 28th May 2008]

Neubauer W., Doneus M., Studnicka N., Riegl J., 2005. Combined high resolution laser scanning and photogrammetrical documentation of the pyramids at Giza. *CIPA XX International Symposium, Torino, Italy* 6 pages. http://cipa.icomos.org/fileadmin/papers/Torino2005/470.pdf [Accesed: 28th May 2008]

Ou Yang D., Feng H-Y., 2005. On the normal vector estimation for point cloud data from smooth surfaces. *Computer-Aided Design 37*: 1071-1079.

Ressl C., Haring A., Briese C., Rottensteiner F., 2006. A concept for adaptative mono-plotting using images and laserscanner data. *ISPRS Commission III Symposium "PCV'06 Photogrammetric Computer Vision", Bonn, Germany*. pages: 98-104. http://www.isprs.org/commission3/proceedings06/complete_proceedings/final-proceeding.pdf [Accesed: 28th May 2008]

Reulke R., 2006. Combination of distance data with high resolution images. *ISPRS Commission V Symposium "Image Engineering and Vision Metrology", Dresden, Germany*. 5 pages. http://www.isprs.org/commission5/proceedings06/paper/1277_Dresden06.pdf [Accesed: 28th May 2008]

Schenk, T., 2000. Object recognition in digital photogrammetry. *Photogrammetric Record*, 16(95): 743–762.

Schumann l., Schoenwaelder B., 2002. Color codes stereo vision in VRML. http://www-vrl.umich.edu/sel_prj/chroma3d/ [Accesed: 28th May 2008]

Yokoyama H., Chikatsu H., 2006. Automatic break-line detection from laser scanner data using surface flatness. *ISPRS Commission V Symposium "Image Engineering and Vision Metrology", Dresden, Germany*. 6 pages. http://www.isprs.org/commission5/proceedings06/paper/1251_Dresden06.pdf [Accesed: 28th May 2008]

9. ACKNOWLEDGEMENTS

The authors would like to clarify that they bear no relation to the owner of the Chromadepth® system, American Paper Optics, Inc. and that this paper reports an independent implementation of their technology.

InMan: How to make sustainable visualisations of the past

D. Pletinckx

Visual Dimension bvba, B-9700 Ename, Belgium

KEY WORDS: historical reconstruction, 3D visualisation, EPOCH, virtual heritage, London Charter, long term availability

ABSTRACT:
Current technology allows us to more and more easily create three-dimensional models of landscapes and man-made structures and to visualise these models in several interactive and non-interactive ways.

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, called InMan, 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.

1. INTRODUCTION

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.

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.

3D visualisation however had and partially still has the connotation of lacking crediblity and documentation, 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 [RYA01]. 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].

We avoid the term *virtual reconstruction* because our main goal is not to reconstruct the past – this is something we simply cannot 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 in some cases we avoid 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. WHY INTERPRETATION MANAGEMENT?

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" a 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 InMan 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 InMan methodology is 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 also 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 physical 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 see this within its context 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 are not that lucky 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 be finished. 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, etc.) 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 is part of the QuickTime software that is able to visualise panoramical and spherical images and interactive objects.

Such 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 truth 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.

4. STRUCTURE OF THE INMAN 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.

A detailed description of the structure and functionality of the InMan tool and the advantages of its Wiki implementation can be found in [PLE08].

5. 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 lack 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 has no 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.

6. BENEFITS

The InMan 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.

7. 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.

The InMan 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 thank 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 also 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.

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 Jun 1, 2008

[EPOCH] EPOCH Network of Excellence, http://www.epoch-net.org/, last verified on Jun 1, 2008

[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 Jun 1, 2008

[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 Jun 1, 2008

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

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

[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 Jun 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

[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 Jun 1, 2008

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

TRACEABLE STORAGE AND TRANSMISSION OF 3D COLOUR SCAN DATA SETS

Stuart Robson [a] , Ian Brown [b], Mona Hess [*] [c], Sally MacDonald [c], Yean-Hoon Ong [c] , Francesca Simon Millar [d] , Graeme Were [d].

[a] Department of Civil, Environmental and Geomatic Engineering, University College London – srobson@cege.ucl.ac.uk
[b] Oxford Internet Institute, Oxford University – ian.brown@oii.ox.ac.uk
[c] UCL Museums and Collections – m.hess@ucl.ac.uk, s.macdonald@ucl.ac.uk, yean-hoon.ong@ucl.ac.uk
[d] Department of Anthropology, University College London – f.millar@ucl.ac.uk, g.were@ucl.ac.uk
University College London, Gower Street, London, WC1E 6BT, United Kingdom
e-curator@ucl.ac.uk

KEYWORDS: 3D colour laser scanning, traceability of data collection, digital documentation, metadata, e-Science, secure data transmission

ABSTRACT:

This paper describes a practical multidisciplinary approach to the traceable storage and transmission of 3D colour scan datasets using a combination of state of the art colour laser scanning technology, an e-Science developed data storage and retrieval solution (SRB) and an internet capable 3D visualisation tool which is being iteratively designed in association with a team of museum curators and conservators who are able to directly compare the handling of a range of original objects with their virtual copies.
Examples taken from a core object group from the world-class UCL Museums and Collections that have been recorded with a state of the art Arius3D colour laser scanner provide a demonstration of the developed 3D recording methodology and highlight how the developed system is capable of complementing traditional cataloguing methods for museum artefacts.

1. INTRODUCTION - THE UCL E-CURATOR PROJECT

The E-Curator project commenced in October 2007 at University College London (UCL) and is an interdisciplinary project which draws on UCL's expertise both in curatorship, in e-Science and the application of laser scanning technologies. Realising the importance of digital technologies and new interdisciplinary possibilities, the E-curator project is being undertaken by UCL Museums and Collections with the goal of applying two state of the art digital technologies: 3D colour laser scanning and e-Science.
This project captures and shares very large three-dimensional scans and detailed datasets of a variety of museum artefacts in a secure computing environment. The E-curator project presents an opportunity to exploit and analyse e-Science technologies and explore some of the opportunities they offer to museum practice in an increasingly virtual world. Such deployment will enable curators and conservators to compare records collected at different institutions, stored remotely, or collected over a period of time under different conditions, in order to assess and monitor change. Our complementary research aims to provide a useful user-friendly web-based shared platform for curators, conservators, heritage and museum specialists.

2. TRACEABLE COLLECTION OF 3D DATA SETS

A first aspect of our transdiciplinary approach is in the traceable collection, storage and sharing of data. The need for traceability (Cooper/Robson 1994), often termed provenance in the fine art domain, has necessitated the development of a general structure for scanned 3D colour data sets that can encompass not only the properties of the scanning system used, but also the chain of processing steps needed to develop a hierarchy of 3D data sets appropriate for viewing and analysis. We will detail the development of such scanning and post processing procedures in combination with a compatible set of metadata standards for 3D laser scan data. This area has been developed much more broadly in the form of the London Charter (Beacham/ Niccolucci 2006) to encompass themes of integrity, transparency, quality and community amongst a broad definition of interest of research and communication in cultural heritage.

2.1 Properties of the scanning system

The 'Arius3D Foundation Model 150' scanner (see Figure 1 and web reference 2) offers a detailed non-contact and non-destructive documentation and examination method which predetermines its use for conservation recording.
Surface scanning is carried out by a scanner head which emits a laser beam composed of three discrete red, green and blue wavelengths, which is focussed to deliver a spot diameter of the order of 100 micrometers within a 50mm deep scanning field. The scan head simultaneously measures surface reflectance at each of the three wavelengths and geometry by triangulation to record the laser reflection at each illuminated location. Every point for which a surface reflection is sensed by the scanning system has therefore a XYZ coordinate location and an RGB colour value. Given the illumination and collection angles at each point as well as at neighbouring points it is possible to also compute a surface normal at each illuminated point. The optical arrangement used is based on the National Research Council's patented auto-synchronized spot scanning principle (Taylor et al. 2003). A calibrated 'white cube' made of diffusely reflecting "Spectralon" (see web reference 4) is illuminated at the end of each scan line to provide a consistent reference surface enabling fluctuations in background illumination and laser output power to be corrected on a scan line by scan line basis. This correction is carried out as part of an off-line colour calibration process in combination with the scanning of a grey scale step series and planar and spherical Spectralon reference surfaces.
The Arius3D scanner head is mounted to a coordinate measuring machine which provides both the dimensional stability and the metrology level motion control necessary to repeatedly move the scan head over the object to be recorded

along 3D paths determined by the user. The depth measurement capability of the scanning head in use at UCL has been proven to be better than 25 microns in depth following the scanning of several reference objects scanned for a range of engineering projects, for example in working with the UK Atomic Energy Authority and EFTA JET (Brownhill et al. 2007). The motion control system has been tested according to ISO 10360-2 (ISO 10360-2:1994) and is designed to deliver a minimum spatial sampling interval of 100 microns, commensurate with the laser spot diameter used to sample the surface within the field of view of the scanning head. To ensure consistent dimensional capability the unit has been installed in an air-conditioned room that maintains temperature at 20°C and can control relative humidity over a wide range to suit the requirements of the objects being recorded.

2.2 Scanning methods

Objects to be scanned with the system are supported either on a turntable or other rigid structure on the flat table beneath the bridge of the motion control unit. The scan head is then guided over the object by the operator in order to check appropriate surface reflection, assess best orientation between the scan head and the object for efficient scanning and to make sure that the combined RGB laser beam has an appropriate intensity to record the object surface. As with all scanning techniques, successful surface recording is heavily dependent on the optical reflectance properties of each individual object with diffusely reflecting surfaces, such as terracotta being relatively easy to scan whilst highly specular surfaces or very dark surfaces can severely test the capabilities of both the operator and capture technology to provide data that are fit for purpose.
Once appropriate scanner settings and an appreciation of the object properties and geometry have been gained by the operator scanning commences with the scan head being driven across the object surface, typically using both X (bridge) and Z (depth) driving systems of the scanner. Colour scan data are collected into Arius A3DScan software with data being collected in swathes 50mm wide at roughly 250mm per minute when scanning at the highest 100 micrometer point density. The A3DScan software environment allows the operator to visualise data, to assess overlap between component scans and to monitor live profile and colour data for example. On completing each group of scans the object or the scan head are manipulated to present a next set of surfaces to the user. At UCL object manipulation is carried out either by a curator or an approved and trained object handling specialist. Scan data collected during the primary recording process are archived to provide a raw data record before any further data processing is carried out.

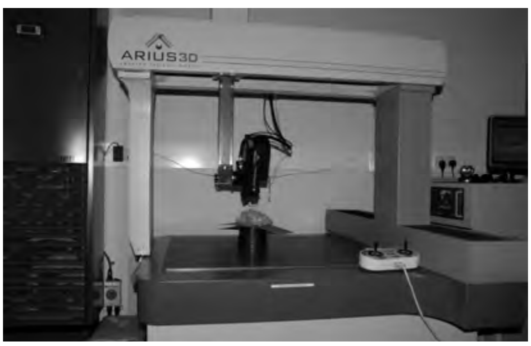


Figure 1: The Arius 3D scanner installed at UCL

2.3 Post-processing of 3D colour laser scan data

Before the scan data is ready for dissemination to an end user it must be corrected for colour recording differences due to imaging geometry and surface shape, registered to take into account arbitrary movements of the object and then cleaned and edited to provide consistent point spacing and colour. As a first step the raw scan data is imported into the proprietary Arius3D, Pointstream 3D Image Suite software (see web reference 3), which is specifically designed to process and visualize the densely sampled 'point cloud' data. The import function allows the user to set options for data smoothing, apply a geometry based colour correction based on data acquired from scans of colour and geometry reference objects and to optionally remove isolated data. The process also computes a normal vector per point to allow application of computer lighting models without the need to convert the data into polygons. This particular aspect takes best advantage of the very dense data produced by the system to visualise using point rendering graphics (Grossmann/Dally 1998) rather than having the overhead of generating polygonal meshes that are common in most other approaches.
The subsequent registration step controls the optimum alignment of the different component scan records in order to register them together using a 'point cloud to point cloud' iterative closest point procedure founded on the method described by Besel and MacKay (Besel/Mckay 1992). The point cloud data are then cleaned within the Pointstream application to remove overlapping geometry and balance colour acquired from different scanner head viewpoints using both manual editing and automatic filtering of multiple points based on combinations of geometry and colour content. After editing the geometry and colour information the component records are merged into a single scan data set.

2.4 File and metadata handling within E-Curator

Every three-dimensional record is annotated with metadata concerning its capture, the import and colour calibration filters used, and any particular scan process and post processing information (see
Table 1). This dataset provides the user with clear information about the production and the authorship of the 3D image.
The development of the metadata set has largely followed the recommendations of English Heritage for scan data (English Heritage 2006) and 'Big Data' project for Arts and Humanities (ADS and EH 2007). The metadata can be displayed alongside the 3D colour model in the prototype E-Curator application (see Figure 11).
For E-Curator these metadata guidelines have been extended to accommodate different sets of raw and processed data in order to provide traceability and transparency for the user, who is interested in examining the 3D images generated at different scanning and data processing stages. A hierarchy for these records has been established including the aligned 'registered' version of the point cloud without colour or point processing, a 'processed' version with cleaned colours and geometry, and a 'presentation' file with optimised colours and filled data voids. For the use of conservators and curators the second 'processed' model will be the most relevant since it encompasses the complete object in one data set but has undergone the least data alteration and processing.

Object information
Object_Dimension, Object_ID, Object_Descriptor, Object_LocationinMuseum, Object_LocationFound, Object_Period, Object_Owner, Object_Collector, Object_Description, Object_ModernLabel, Object_Inscription, Object_Material, Object_Category, Object_Name, Object_OriginalClassification, Object_Provenance, Object_ResearchQuestion, Object_LinktoWebDatabase
File information
File_NumberofPoints, File_ImportSettings, File_NumberofRecords
Scan information
Project_ID, Scan_Name, Scan_Hardware_used, Scan_Software_used, Scan_LaserSettings, Scan_Scandate, Scan_Time, Scan_Scanagent, Scan_Comments, Scan_Description, Scan_ColourCalibrationSet
Processing information
Process_ID, Process_Agent, Process_Software_used, Process_DateStartProcessing, Process_DateCompletionProcessing, Process_Time Process_Description, Process_Guidelines
Database information
Database_Added (to database), Database_Added by

Table 1: Metadata information for a single scan in the E-Curator prototype (July 2008)

It is also necessary to consider the technical challenges posed in maintaining quality in colour and geometry data collected with scanning systems, both of which must be carefully controlled to ensure that the collected record is fit for virtual heritage purposes. Whilst specifications for recording the position, size and departure from nominal form of geometric objects exist in the engineering domain (BS 7172: 1989) recording the geometry of free form surfaces to a similar level of best practice is in its infancy.
The situation is more complex for the recording of colour where knowledge and best practice is widely distributed amongst different user communities ranging from paint manufacturers through the digital imaging domain to surface finish specialisations. Most pertinent are models of Bidirectional Reflectance Distribution Function (BRDF) which can account for the variation in directional dependence of the reflected energy from a surface. Calibrated reflectance test charts, such as those in the MacBeth series, can of course be included in the process and are able to account for capture and viewing characteristics, but they are unable to correctly represent the surface characteristics of individual object to be recorded.
In the case of the Arius3D the system colour collection is reproducibly monitored for each scan pass and many of the key effects inherent in the imaging configuration are corrected, but there is no account made of viewing angle dependant colour surface properties. As a result the data produced following the editing process can be regarded as an operator's best attempt at a consistent and engaging 3D colour model rather than being consistent with an external colour standard.

3. EXAMPLES FOR 3D COLOUR LASER SCANNING FROM UCL MUSEUMS AND COLLECTIONS

A systematic approach and methodology for organic and inorganic materials is being developed within the project to ensure repeatability of the scan process. The resulting coloured point cloud data available in the E-Curator application provide a highly accurate surface analysis tool.
As illustrative examples we would like to introduce a selection of outcomes from the E-Curator core object group, all from the world-class UCL Museums and Collections. The diversity of these collections, from archaeology, art and ethnographic to medical, provide an excellent 'laboratory' to apply and test the recording technique. All examples have been digitized with the Arius3D colour laser scanner at UCL.

3.1 Example 1 – a painting from UCL Art Collections

We chose a very rare example of British Impressionism by outdoor painter Walter Westley Russel. The 'Beach Scene' (ca. 1890) is part of the Paintings Collection in UCL Art Collections.

Figure 2: Three-dimensional colour laser scan of Russel 'Beach Scene' painting. Brush strokes and surface features are clearly visible due to adjusted artificial light settings.

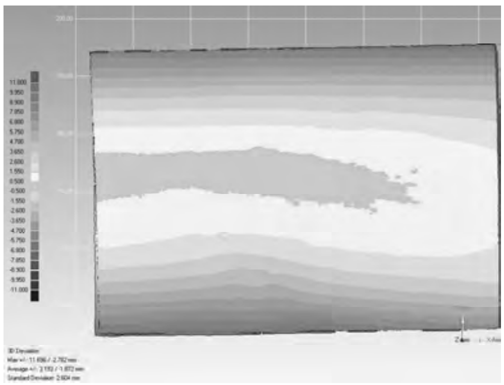


Figure 3: Geometric evaluation of the Russel painting. The convex deformation is now quantifiable from the 3D model.

The lightweight wood used as a sketching panel, possibly deal or mahogany, played an important role in the painting process. Not only did the paint dry rapidly on the unprepared surface but the painting also profits from a warm tonality within its coloration (Strang Print Room UCL 2007). The painting shows some serious conservation problems.
Detailed conservation science analyses (X-ray, pigment analysis) have been done on the painting by the UCL Art History Department and the 3D colour laser scan hopes to complement the results.
The goal is to produce a virtual three dimensional archival document that widens our knowledge of the object.

The scan product is a coloured point cloud with the density of 100 µm per 100 µm; such a detailed survey of the painting has not been delivered before. Brush size, paint direction and structural surface anomalies can be visualised by using the options of artificial raking light settings. The 3D scan, stripped of its colour information, makes surface features more visible to analysis (see
Figure 2 and Figure 4). The now perceptible impressionistic paint structure could lead to knowledge about an artist's personal style - his 'handwriting' - and could help to identify authenticity of a painting.
The 3D colour scan is not only documenting the brushwork of the painting. But also the warping deformation of the wooden panel is recorded, which provides an ideal starting point for detailed surface measurements and further monitoring over time. Sections and evaluations of geometric deformation can easily be derived from the 3D point cloud (see
Figure **3**).

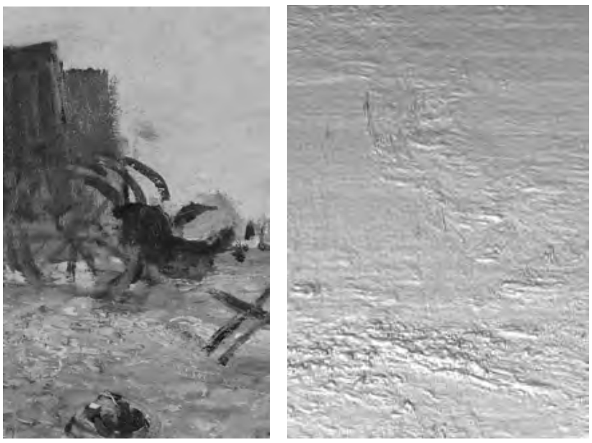


Figure 4: Photographic detail versus 3D laser scan: detail of the Russel painting. The X-ray of the painting gave evidence of earlier additions, maybe more 'bathing machines', in the background. They were later painted out with lead white. A fact that can be also be confirmed by the scan image.

The results suggest that the technique can complement traditional conservation analyses methods.

3.2 Example 2 – a mask from UCL Ethnographic Collection

UCL's broad range of collections includes a Sepik Yam Mask from Papua New Guinea. This fragile mask is woven from natural fibres and decorated with coloured pigments, shell valuables and some feathers. A 3D scan should reveal greater surface detail of the mask, pigmentation loss and deterioration of the woven fibres.
The fragility of the mask prohibits a too frequent lifting and turning upside down. Students of conservation and anthropology could now browse the digital object to find out more about the weaving pattern of the mask without increasing its deterioration by handling.

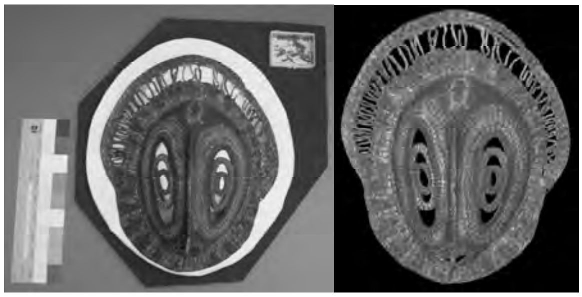


Figure 5: Sepik Yam Mask: photograph versus 3D colour scan.

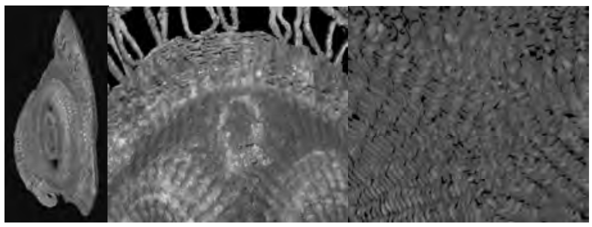


Figure 6: Details of side, front and back of the New Guinea Mask. The front detail is showing pigment and surface loss, whereas the back detail reveals an intricate weaving pattern.

Further distribution methods of this digital model to enhance knowledge about this type of object can be considered now: virtual reconstructions, animations and films for multimedia and education (e.g. in e-learning).

3.3 Example 3 – The Petrie Quartzite: an incised and inscribed stone from 1500 BCE

This object from UCL Petrie Museum is described as an irregular carved slab of quartzite, orange-brown colour, with black-outlined incised cartouche of throne-name of Hatshepsut, Maatkara on the convex side. A line of largely effaced hieratic in black colour runs along the left side of rectangle framing of the cartouche. It is dated into the period of dynasty 18 (1295BCE-1550BCE).
The Petrie Quartzite is frequently loaned but has recently failed a conservation test. The curators hope that a digital image of the stone helps distinguish tool marks from erosion without direct contact with the object surface and facilitate a better reading of the fading carbon-black cursive inscription line.

Figure 7: A catalogue photograph next to a 3D colour laser scan of the namestone aka Petrie Quartzite, UCL Petrie Museum collection number UC 55606.

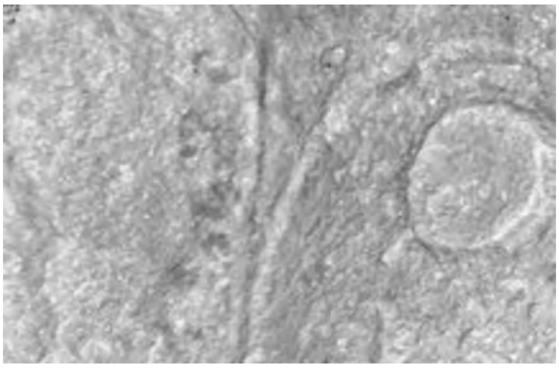


Figure 8: A detail of the colour laser scan which enables the conservator and curator to see fine ink writing as well as surface geometry, geological formation and coloration of the stone.

A 3D colour laser scan documents the object in the round in high resolution. It also reveals almost indiscernible surface incisions and features on this namestone (see Figure **9**).

Further research on the digital image is now possible without endangering the surface by abrasion through manual handling. External researchers could be invited to give their comments remotely and compare it to similar objects.

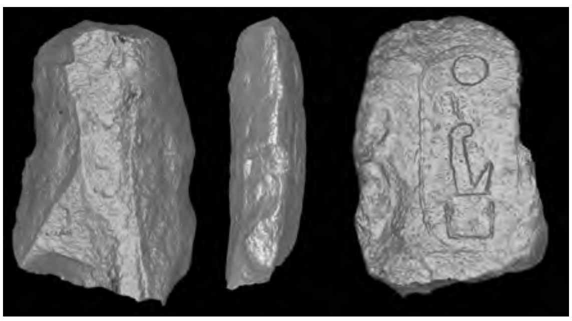


Figure 9: 3D laser scans of the Petrie Quartzite 55606.

4. TRANSMISSION OF 3D DATA SET

A facility for the inter-institutional exchange of the collected data scans and related metadata is being developed to provide a shared platform accessible by museum specialists on low cost computing hardware. The E-Curator application prototype is an e-Science developed data storage and retrieval solution and an internet capable 3D visualisation tool. It is being iteratively designed in association with a team of museum curators and conservators who are able to directly compare the handling of a range of original objects with their virtual copies.

4.1 Data sharing and real time interaction with 3D colour laser scan data

For 3D colour scans to be of practical use, robust means of sharing and validating the data obtained need to be established. High resolution colour scans of one object can require hundreds of megabytes of storage space, and can only realistically be shared using the distributed file systems such as Storage Resource Broker (SRB) being widely deployed in the e-Science environment.

SRB is a data grid middleware software system produced by the San Diego Supercomputer Centre (SDSC). The system implements logical namespaces (distinct from physical file names) and maintains metadata on data-objects (files), users, groups, resources, collections, and other items in an SRB Metadata Catalogue (MCAT), which is stored in a relational database management system. The SRB has features to support the management and collaborative (and controlled) sharing, publication, replication, transfer, and preservation of distributed data collections.

The E-Curator prototype currently enables an external user to access and search different objects across UCL Museums and Collections. The following categories are displayed and can be browsed: Catalogue Information, Conservation Information, Exhibitions & Displays, References, 2D images, related objects in the E-Curator application and 3D images with its scan metadata information.

SRB allows data to be replicated from remote into local databases to allow fast access. Further development will automatically connect E-Curator users to such local versions when available, reducing the time taken to retrieve commonly-accessed objects and also minimising bandwidth demands on remote connections.

4.2 Software architecture of the E-Curator application: e-Science technology (SRB)

The E-Curator application is middleware software that provides user access via a web browser to 3D colour scan data and relevant catalogue and metadata information. It is written as Java servlets running on a Tomcat server (see web reference 1), and interacts with the Storage Resource Broker using the JARGON library (see web reference 5). Users access this software using their web browser (see Figure 10).

Metadata about the scanning process is stored in SRB alongside each scan. Cataloguing information about each object is stored using SRB's Metadata Catalogue. This information is displayed alongside object scans within Internet Explorer (see Figure 11).

3D scans of objects in Arius3D Pointstream format are stored on the SRB server as a collection hierarchy. A web-based interface has been developed to allow users to access the 3D images, using an Arius ActiveX plugin to visualise and manipulate scans in Internet Explorer (see Figure 12).

Institutions wishing to provide access to scans of their own objects will have the choice of installing this software for themselves (which is being released under an open source licence), or alternatively partnering with an organisation that provides this type of functionality – such as in the UK the National Grid Service funded by the UK Research Councils.

Figure 10: E-curator application prototype (June 2008).

Figure 11: E-Curator application prototype while browsing the metadata of the object and 3D scans (June 2008).

Figure 12: Browsing the 3D model of a Medicine Bundle from West Africa (UCL Ethnographic Collection) in the E-Curator application.

4.3 Metadata and catalogue information for 3D scan data

3D representations of heritage objects and museum artefacts need to have a clear 'provenance'. This includes both a description of history and ownership of the object and a clear set of data that describes the production of the digital 3D scans of the object.

The E-Curator team has developed a specific metadata set for the E-Curator application that includes UCL Museums and Collections catalogue entries. The metadata used to describe the 3D images is based on SPECTRUM (1996), the UK Museum Documentation Standard for catalogue entries. These metadata include information about the object ID, physical description, location, historical facts, condition, exhibition and conservation information. The metadata provides a relevant framing for artefacts such as historical and archaeological facts, conservation information, exhibition and display information. The data is collected from local databases and paper documentation currently kept at UCL museums.

5. CONCLUSIONS

The examples given in this paper provide a demonstration of the developed 3D recording methodology and highlight how the developed system is capable of complementing traditional cataloguing methods for museum artefacts. The availability of complete data sets from raw scan data to final edited models allows users to interrogate the available 3D content from a range of standpoints. On a basic level simply to gain an appreciation of the overall appearance of an object from its final edited record and for a more scientific enquiry to access constituent data at a range of defined levels that reflect the data processing regime developed for the Arius 3D scanning system deployed within the project. The hierarchical system has the capability to be expanded to accommodate data from other scanning systems, photogrammetric imaging and Polynomial Texture Mapping (PTM) which are also deployed at UCL.

The utilisation of Grid technology through SRB and an allied web based interface is expected to prove extremely useful in providing a scalable solution for the dissemination of information from the E-Curator project to a multitude of users. Currently the user group is limited to a museums and collections community who are involved with project assessment and the development of further E-Curator capabilities, but the system will be made available to external users on both real and virtual visits to UCL Museums and Collections sites in the near future.

6. BIBLIOGRAPHY AND REFERENCES

Archaeology Data Service and English Heritage (ADS and EH) 2007: Preservation and Management Strategies for Exceptionally Large Data Formats: 'Big Data' (EH Project No. 3948), Final Version 1.03, 28. September 2007. *http://ads.ahds.ac.uk/project/bigdata/* (accessed on 18 May 2008).

Beacham, R., Niccolucci, F., 2006. The London Charter for the use of 3D visualisation in the research and communication of cultural heritage. Draft 1.1 from 14th June 2006 *www.londoncharter.org/downloads.html* (accessed on 18 May 2008).

Besel, P., Mckay, N., 1992. A method for registration of 3D-shapes. *IEEE transactions. Pattern analysis and Machine intelligence*, 14(2): pp. 239-256.

Brownhill, A., Brade, B., Robson, S., 2007. Non-contact surface measurement in a hazardous environment. 8th Conference on Optical 3-D Measurement Techniques, *Proceedings of the 8th Conference on Optical 3-D Measurement Techniques series.* Zurich, Switzerland: Swiss Federal Institute of Technology, Zurich, pp. 64-71.

BS 7172:1989 British Standard Guide to Assessment of position, size and departure from nominal form of geometric features. British Standards OnLine: *http://products.ihs.com/bs-seo/gbm24_15.htm* (accessed on 18 May 2008).

Cooper, M. A. R., Robson, S., 1994. A hierarchy of photogrammetric records for archaeology and architectural history. *ISPRS Journal of Photogrammetry and Remote Sensing* 49(5), pp. 31-37.

English Heritage 2006: An Addendum to the Metric Survey Specifications for English Heritage. Addendum 2: A standard specification for the collection of point cloud data by terrestrial laser scanning or other methods. *www.english-heritage.org.uk/upload/pdf/metric_extraction_scanning_addendum_1.pdf* (accessed on 18 May 2008).

Grossman, J. P., Dally, W. J., 1998: Point sample rendering. In *Rendering Techniques '98*, pp. 181–192, July 1998.

International Standard ISO 10360-2:1994 Coordinate Metrology—Part 2: Performance assessment of coordinate measuring machines. *www.iso.ch.*

Rioux, M., 1994. Digital 3-D imaging: theory and applications. *SPIE Proceedings, Videometrics III*, Boston, MA. October 31-November 4, Vol. 2350, 1994. pp. 2-15.

Strang Print Room UCL 2007: *British Painters on the Beach.* Polly Miller and Anna Cooper. in *From Idea to Object. Painting Practices Revealed.* HAMS Exhibition Catalogue. Strang Print Room, UCL, Autumn 2007. Edited by Libby Sheldon and Noelle Streeton. Pp. 50 – 55.

Taylor, J., Beraldin, J.-A., Godin, G., Cournoyer, L., Baribeau, R., 2003. NRC 3D Technology for Museum and Heritage Applications. *The Journal of Visualization and Computer Animation.* Volume 14, (3), 2003. pp. 121-138.

References from websites:

1) Apache Tomcat version 6.0.16, *www.tomcat.apache.org*

2) Arius 3D Inc., Ontario, Canada, *www.arius3d.com* and *www.arius3d.com/main.html?contentId=1.*

3) Pointstream 3D Image Suite software, Version 2.9.0.0., *www.pointstream.net.*

4) Optical-Grade Spectralon material, Labsphere, Inc. *www.labsphere.com/data/userFiles/Optical%20Grade%20Spectralon%20datasheet_1.pdf.*

5) SDSC Storage Resource Broker and JARGON library, *www.sdsc.edu/srb/index.php/Main_Page*

6) UCL Museums and Collections - E-Curator website: *www.museums.ucl.ac.uk/research/ecurator/.*

7. ACKNOWLEDGEMENTS

The project is jointly funded by the Arts and Humanities Research Council (AHRC), the Engineering and Physical Sciences Research Council (EPSRC) and the Joint Information Systems Committee (JISC).
We would like to thank the UCL Museums and Collections for the loan and handling of their objects for scanning with the Arius 3D scanner at the UCL Chorley Institute – thanks especially to Andrea Frederickson and Subahtra Das - UCL Art Collections; Ivor Pridden, Stephen Quirke, Susanna Pancaldo - UCL Petrie Museum; Sarah Mengler – UCL Ethnographic Collections; Mark Carnall – UCL Grant Museum of Zoology.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis Full Papers

SURVEYING, MODELING AND COMMUNICATION TECHNIQUES FOR THE DOCUMENTATION OF MEDIEVAL WOODEN PAINTED CEILINGS IN THE MEDITERRANEAN AREA

F. Agnello [a], M. Cannella [a], A. Gentile [b], M. Lo Brutto [a], A. Santangelo [b], B. Villa [a, *]

[a] Dipartimento di Rappresentazione, University of Palermo, Via Cavour 118, Palermo, Italy
fabrizio.agnello@unipa.it, lobrutto@unipa.it, bevilla@unipa.it
[b] Dipartimento di Ingegneria Informatica, University of Palermo, Viale delle Scienze, Palermo, Italy
gentile@unipa.it, santangelo.antonella@gmail.com

KEY WORDS: Cultural Heritage, Survey, Visualization, Multimodal guides.

ABSTRACT:

Wooden painted ceilings of the Mediterranean area in the middle age have their origin in the islamic culture and were then spread in the countries under the dominion of the Arabs; some of the surviving ceilings are now located in Sicily and Spain. In the historic centre of Palermo two well preserved medieval ceilings are still surviving; the first, built in the XII century, is located in the Palatine chapel; the second one, built in the XIV century covers the "Sala Magna" in the Steri of Palermo.
The research, focused on the ceiling in the Steri, deals with the definition of a process for the integration of surveying techniques (photogrammetry, laser scanning), modelling processes and communication technologies for the documentation of such artefacts.
The documentation of painted ceilings requires the strict integration of photographic and 3D metric data; the existing documentation is usually made of documents (drawings, photographs) that keep geometric and metric data separated from the photographic documentation of the paintings.
The first stage in this work is therefore addressed to produce a digital document that combines metric and photographic data in a 3D textured model; in the second stage a vocal guide interacting with the 3D model has been developed; such guide, thought as a support to people visiting the Steri, uses a database with historic contents and symbolic interpretation of the painted scenes to answer specific questions and "take" the visitor close to the related paintings.

1.INTRODUCTION

The high level of automation achieved in the survey operations and the possibility of managing, through the use of dedicated software, a lot of information offer today new opportunities within the field of conservation and communication of Cultural Heritage but at the same time have introduced new problems. An important aspect, object of great interest by the international scientific community, concerns with the communication between the responsibles for data acquisition, elaboration and management (surveyors, photogrammetrists, GIS operators) and the users of information (archaeologists, historians, experts in the field of restoration).
The paper deals with the study of the problems related to the survey and communication of Cultural Heritage through the various phases, from data acquisition and processing to data management. The experiences gained in this field have always shown more and more attention of the experts and historians towards new technologies; however the access difficulty to information often interferes with this interest. Today the automatic acquisition of geometrical data, the attribution of space data and the management of all information concerning an artefact by a single tool, are indispensable components both for data documentation and communication.
The research work is organized into two stages. In the first stage different techniques of survey have been tested and the problems related to the integration of data acquired have been studied. The definition of an homogeneous datum able to describe the geometrical and qualitative complexity of the chosen artefacts represents one of the principal results to achieve. For this purpose, it is necessary to define clear operating methodologies both for raw data acquisition and processing. The main goal of this stage is to come to a real and complete 3D data acquisition. This result can be achieved using laser scanning and digital photogrammetric methodologies. Within these techniques the methodological approaches can be different. In the specific case, a long range laser has been used to realize a general 3D model of the artefact. The realization of high resolution textures mapping is an interesting field of this work. The problem can be tackled in two different ways: texturing the surface models obtained from the clouds of points or linking a corresponding RGB value to each acquired point. The quality of colour acquisition carried out by the 3D laser scanners is in most cases poor because the integrated sensors are characterized by a low geometric resolution. The implementation of suitable algorithms, in order to associate the colour to clouds points from the high resolution images, represents a good solution. In the last years the photogrammetric multi-imagine techniques have had a quick diffusion not only because they are more flexible then the classical stereoscopic techniques but above all for the possibility of realizing three-dimensional surface models from monoscopic images. This approach can be favourable for the realization of models of artefacts characterized by a simple geometry to integrate if possible by high resolution laser scanning techniques. The second stage is addressed to test

* Corresponding author.

advanced methodologies for management and communication of cultural heritage. The research work has been directed to define a standard model in querying the model through a multimodal guide. The main goal of this stage is to determine a procedure in processing 3D textured models so to make them become the base structure for interactive databases.

2.THE WOODEN CEILING IN THE SALA MAGNA

The Steri Palace in Palermo, residence of the Chiaramonte family, was built starting from 1320 on the eastern edge of Piazza Marina, near the city's ancient harbor. From 1605 to 1782, the Palace was the headquarters of the Inquisition Tribunal, a place of detention and torture. Since 1984 the Steri has been the headquarter of the Rector's Office of Palermo University, that has commissioned the restoration works which have given the building its current feature. The "Sala Magna" is a hall of rectangular shape (Figure 1), sited at the Northern side of the building's first level and was the most important room of the palace, used for public events.

Figure 1. The "Sala Magna"

The wooden ceiling covering the hall is made of twenty-four beams laid transversally, and lacunars covering the empty space between the beams. The beams are fitted to the walls and laid on consoles (Figure 2). Scenes related to the same subject are developed on vertical beam faces, all heading to the same direction, in a sequence that goes, according to the point of view, from left to right, and from the background to the foreground.

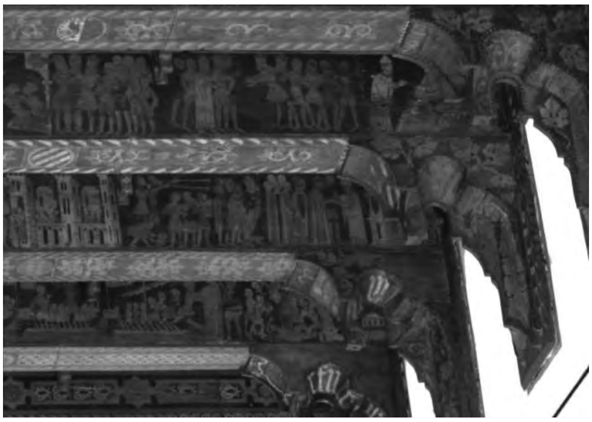


Figure 2. The consoles bearing the beams

The partitioning of the panels of each beam was determined by the articulation of the narration into distinct scenes, according to a technique resembling strip cartoons.
The ceiling has been made in 4 years from 1377 and 1380; it was painted by three local artists that were probably supported by books or by people having a good education in choosing the subjects of the scenes. Historians keep this ceiling as an important document of the middle age cultural references in Sicily: paintings are organized as short tales whose subject are taken from Greek mythology, the Old Testament, and medieval legends (Figure 3). The Chiaromonte family, who had started the construction of the Steri as their residence one century before, represented in this ceiling their cultural ambition as a counterpart of their political role.

Figure 3. Scenes from the paintings. *Above*: Elena and Paride come to Troy. *Below*: Giuditta cuts off Salomone's head

The transformation of the Steri in residence of the Spanish governors, then in Court of the Holy Inquisition and last in public administration office has deeply altered the structure and the distribution of the original building, but has fortunately kept safe the wooden ceiling, hidden by false vaults. The restoration of the Steri, promoted by the university of Palermo that has established here its central administration, has given the ceiling its original aspect.

3. DATA COLLECTION

Metric data were collected with laser scanning devices[1]. Laser scans were used for the acquisition of a remarkable number of 3D points of the ceiling. The scanning operations were conditioned by the complex geometry of the ceiling. In order to reduce the holes corresponding to areas not reached by the laser ray, numerous scans were performed from different points inside the hall. Laser data were collected first with a *Mensi GS200* scanner and in a second stage with a *Faro LS880* scanner (Figure 4).

Figure 4. Laser scanning survey. *Left*: Mensi GS 200 and reflecting targets on the wall. *Right*: Faro LS880

Point clouds collected with Faro laser device have covered, thanks to the wide field of view on the vertical plane, all the areas that could not be reached by the Mensi scanner e.g. the beams and the lacunars near to the walls of the Sala Magna.

Point clouds were oriented and referred using the correspondence between laser and topographical coordinates of targets[2]. Laser scanning survey of the ceiling produced a 15-million-point cloud.

Photogrammetric surveying was aimed at producing rectified images of the vertical and horizontal faces of the beams, and of an orthophoto of the whole ceiling. In fact, considering the shape of the surfaces of the elements of the ceiling (almost flat), there was no need to use bundle adjustment techniques. A *Canon EOS mark II* digital camera, equipped with a full-format 16.1 megapixel CCD sensor and a 50 mm lens, and *Photometric* software by Geotop s.r.l. were used. In the first stage of photogrammetric survey zenith photos of the horizontal faces of the beams and of the lacunars were captured; vertical faces of the beams and consoles were shot with convergent photographs. Photos of the beams and of the lacunars were rectified using the coordinates of an adequate amount of the control points, referred to the mean plane of each element.

4. 3D MODELLING

The 3D model[3] of the hall was developed using laser scanning data; usually 3D surface models are extracted from the point cloud in an automatic way; the result are mesh surfaces whose resolution (density of triangles) is almost uniform. In this work the 3D model was produced in a different way: planar sections and surface interpolation were used to extract from the laser data the features needed in the modelling process (Figure 5).

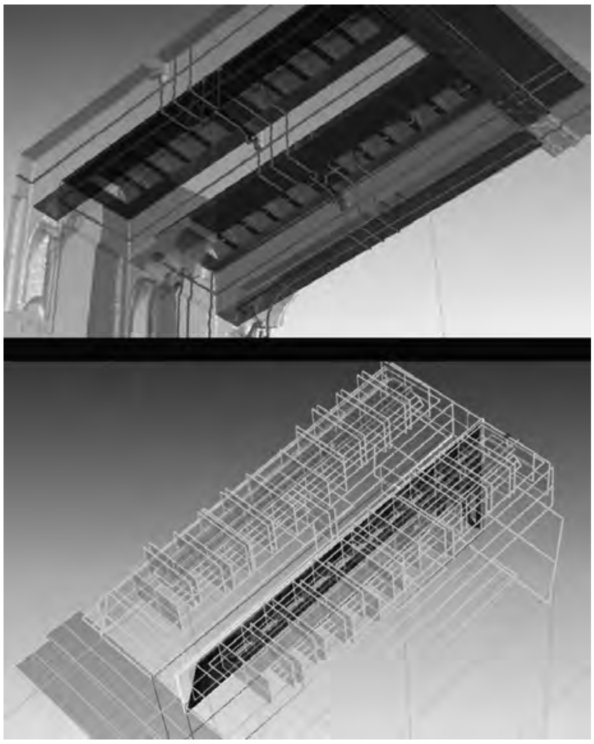


Figure 5. Features extraction from laser data

This way a certain "gap" between the real morphology and the 3D model was produced, but small gaps (deviations never exceed 7 mm) allow us to produce a flexible and handly model of the entire ceiling, with a resolution that is adequate to the local geometry of the surface. It takes to remind that the goal of this work is the production of a 3D interactive model as a support for the documentation of the paintings and as an help for people visiting the Steri. This process, directed to the use of the CAD models for visualization, rendering, or prototyping, requires a reconversion of the digital model into a triangular mesh. The result of this process is a new mesh, where the distribution of the triangles is linked to the shape of the surfaces (Figure 6).

---

[1] The survey was performed in the framework of an agreement between the Regional Institute for the Catalogue and Documentation of Cultural Heritage, the University of Palermo and the Department of Representation. The instruments used were made available by the "Lab for management and enjoyment of cultural heritage with advanced IT" belonging to Palermo University Lab System (UniNetLab).

[2] Laser data processing was performed with Inus Rapidform 2006 and Rapidform XOS.

[3] The 3D model carried out with the package Rhinoceros 4.0.

Figure 6. View of the 3D model

5. VISUALIZATION

In the ceiling covering the "Sala Magna" paintings are not less important than the elements defining structure and morphology. Therefore why the texturing process was cared not less than surveying and modelling. Rectified images were used for mapping beams and lacunars; convergent photos were used for false consoles.

The texturing process, which consists in assigning the raster image UVW coordinates that are linked to the XYZ coordinates of the model, was quite easy for the flat faces of beams, where just a dimensional and position adjustment was needed (Figure 7).

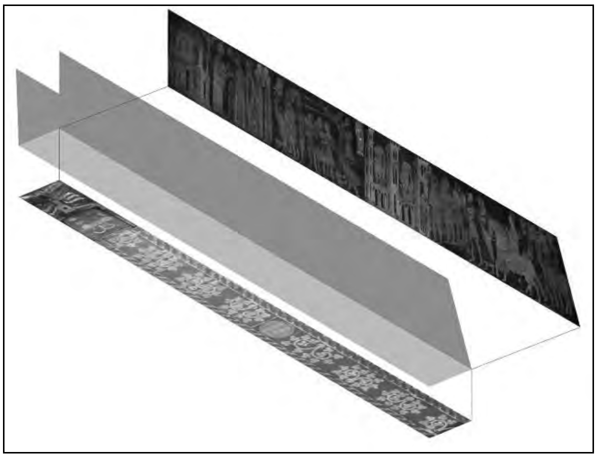


Figure 7. Texturing the faces of a beam

The mapping of false consoles was more complex and harder, since it was necessary to proceed by trial and testing; the quality of this mapping is therefore directly related to the sensitivity and skill of the operator (Figure 8).

The 3-D model thus obtained allows the observation of paintings in their spatial context; it gathers what has always been distinguished in literature, i.e. paintings – described with photos – and the physical support - described by graphic or physical models.

The transformation of the model into a generic format for visualization (VRML) makes 3D exploration accessible also to users who do not have specific education or devices (Figure 9).

Figure 8. Texturing of a console

The VRML model can also be accessed via internet, allowing users from distant countries to "visit" the Sala Magna.

VRML files can be edited in different ways and digital contents can be linked and accessed during the navigation; this way the model becomes something like a 3D homepage of a hypertext.

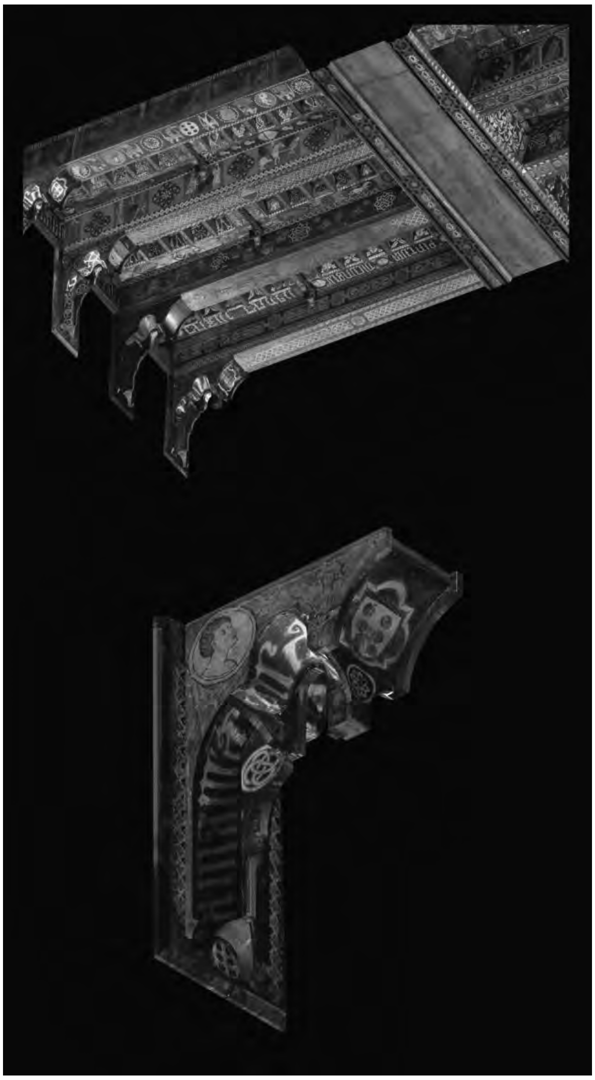


Figure 9. Views of the textured model

The user can freely navigate and inspect the model; if the cursor comes over a sensible area its form changes and a note informs that a link is accessible from that point. Multimedia files (raster images, html pages, music, videos, text) or further 3D models can be linked to the VRML file.

The model becomes an "open structure" whose extension is limited only by the imagination and the work of its creators. In the next future it is not hard to imagine a digital database linked to the 3D model, open to contributions from scientists from different countries.

VRML models can also be used to produce multimodal guides as effective support to people visiting the monument or for people accessing the Sala Magna via internet.

The research has received a meaningful contribution by teachers and researchers in computer science from the university of Palermo; the following section of this paper reports the results of a study addressed to the production of a multimodal guide for the wooden ceiling of the Steri.

6. THE MULTIMODAL GUIDE

The area of cultural heritage preservation and fruition has drawn an ever growing attention of artificial intelligence and human-computer interaction research in the last decades. This interest has led to the creation of clever and clever systems that can interact with the user in a variety of modes and in the most natural way.

The use of virtuality and multimodality in Cultural Heritage fruition allows for a solution to the contrast between the conservative and expositive function of the cultural heritage. This approach gives a trade-off between the need to preserve unchanged the authenticity of the deteriorated heritage and the demand to aesthetically make it comprehensible and enjoyable.

The proposed solution is based on a Multimodal Browser and aims at assisting users during the access to collateral information. The system embed the virtual environment with a multimodal interactive guide which assists the user during his virtual tour. Visitors can navigate the virtual representation of the ancient wood ceiling with tempera paintings and achieve, interacting vocally, relevant meta information about *history and background of painted scenes.*

6.1 X3D and Voice integration

To address this issue one of the chances is the definition of open common standards for interface development. Nowadays integration of the hypertext HTML with VRML (Virtual Reality Modelling Language) has been already standardized providing a good approach in order to augment the efficiency and usability of 3D user interfaces. Starting from this point, in our work, we present integration of X3D (eXtensible 3D), an evolution of the VRML, with XHTML+Voice, an extension of HTML with the VoiceXML, an hypertext language for the voice interaction management.

As said before the interactive virtual environments has been implemented using the eXtensible 3D (X3D) technology. It is a new Open Standard developed by the Web 3D Consortium as evolution of the Virtual Reality Modelling Language, VRML.

In comparison with his predecessor VRML (that has been stopped to the version 2.0), X3D adds new nodes (e.g. the curved NURBSs and the Humanoid Animation).

The basic structure of the X3D environment at run-time is the scenes graph, in which all the objects of the system and their relationships are contained. The X3D run-time environment deals with different functionalities: visualizes the scene, receives input from different sources or sensories, coordinates the events process. Also it manages the current state of the scenes graph, the connections among the X3D browser and the external applications for the hyperlinking, the access through API, the cycle of life of the single objects, both built-in and defined by the consumer.

The X3D files were managed with the X3D-Edit development environment that provides developers with tools for the implementation of interactive 3D scenarios. The X3D standard supplies the possibility to set events inside the 3D worlds, through the use of script code.

In the X3D model of the application, some prearranged events are identified that allow to manage the interaction with the multimodal guide. This connection is managed through the javascript language.

Specifically, inside the 3D environment, some Grouping nodes, called anchors, were implemented each one able to contains different nodes. The Anchor grouping node retrieves the content of a URL when the user activates (e.g., clicks) some geometry contained within the Anchor node's children. If the URL points to a valid X3D file, that world replaces the world of which the Anchor node is a part, if non-X3D data is retrieved, the browser shall determine how to handle that data; typically, it will be passed to an appropriate non-X3D browser.

The implemented system overworks this second features of the anchor node. In our particular instance, the URL in the anchor node contains the code of an ad-hoc javascript function, that activates the vocal interface, as shown in Figure 10. This function loads an XHTML+Voice page that manage the vocal interaction with the user.

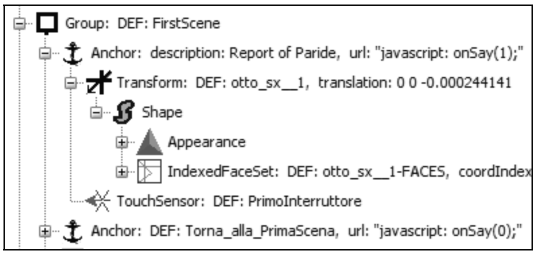


Figure 10. X3D-Edit environment screenshot, Anchor node for the First Scene

The geometries, to which hyperlinks are "anchored" to, are well-defined areas of the implemented 3D environment, that were selected during the designing phase. For each area and ad-hoc Viewpoint node was introduced in the scenes graph, as shown in Figure 11.

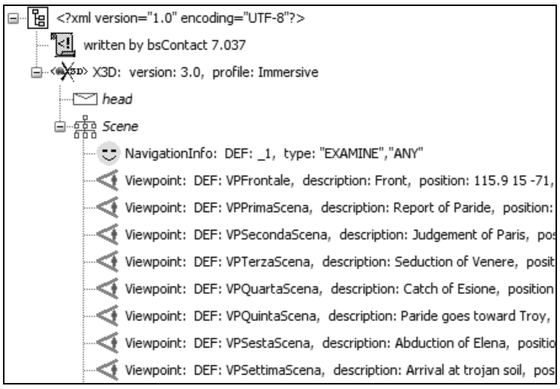


Figure 11. X3D-Edit environment screenshot, Viewpoint nodes

The Viewpoint node is considered as camera point of view that visualize a selected area in the world. When a Viewpoint is "active" the user see the world from this point of view.
The choice of the areas were made taking into consideration two aspects. The relevance of the area inside its context and the called forth interest of the scene for the user experience. The detected zones of interest in the virtual word are then tagged and highlighted with a tooltip that shortly describes the point of interest.
Once selected a point of interest, the vocal guide accomplishes the task to give information about the area to the user. A set of javascript functions manage voice guide activation coherently to the visual content of the page.
In Figure 12 is shown the integration between X+V and X3D technologies.
The interaction is performed with two different modality: visual input are used to navigate the virtual environment; vocal input are used to navigate meta content about the domain supplied by the virtual guide. During navigation the user can freely select the path of his tour, reminding that visual command take priority over vocal command.

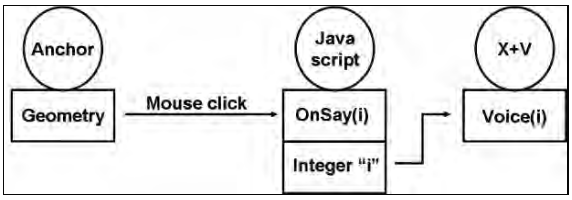


Figure 12. X3D and Voice integration

6.2 The prototype

The implemented prototype aims at assisting a user during a visit in the virtual environment. Furthermore, as the painted scenes follow a logical path according to the story of the Trojan Cycle, a visitor would better enjoy its sight if supported by a detailed description of the scenes and their related background.
To create an appropriate dialogue the information about the context has been extracted and translated in English from historical source material and stored in a database.

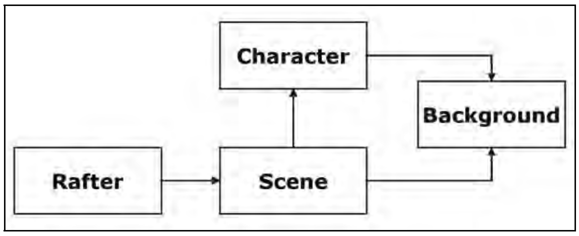


Figure 13. Formalization of context information

As shown in Figure 13, historical information about the paintings were divided in: scene description, character genealogy and background that is referred to a scene as prior events and/or to characters as flashback.
Interface features were decided by evaluating possible visit scenarios, assuming the virtual representation as a way to augment visitor experience on site.
The interaction is centred on the rafters. The user can visit the room and select a specific rafter to look in detail. Specific scenes of the rafter can be visualized, and historical/artistic voice-over be heard at the same time. At any time during the visit, a visitor can change viewpoint, leaving him to explore the virtual environment according to his preferences.
The initial screen of the application represents the entire room. It is possible to navigate inside the room with the functionalities made available by the X3D player. A welcome message introduces the visitor to the availability of a vocal guide, and gives some preliminary information on how to interact with the system.
During navigation, when rafters are in sight, some tooltips, shortly describing the content of each rafter, guide the visitor through the tour. When the user selects a rafter, the point of view changes showing all the scenes in the rafter, while the vocal guide gives some information to the user, as shown in Figure 14. The screenshot shows also the guide window in foreground. At all times, this windows can be hidden, moved or resized, according to the scene elements.

System: You are looking at rafters eight and nine, showing the Trojan Cycle. Click on the scene to zoom in and to listen more.
User: (Click on the scene: "Report of Paride").

Figure 14. Dialogue and screenshot in the rafter perspective.

In this view, the various scenes about the Trojan Cycle are shown to the user and each one can be selected to access more details. Once a scene is chosen, it will be zoomed in, the multimodal guide will be activated, and the historical/artistic information will be delivered. The guide will prompt the visitor to solicit interaction for further details, as described in the sample dialogue in Figure 15, where the visitor may be querying about a scene or access information about its characters. Narration can be skipped at any time.

System: You are looking the scene called Report of Paride. Do you like to listen some detail about this scene?
User: Yes.
System: In this scene Paride narrates the judgment to Priamo and Trojan nobility, which is a preview of a following scene. Are you interested in more details about characters story or background of the scene?
User: Yes, thanks.
System: Which characters are you interested in?
User: Help.
System: Say the name of the character, choosing between Paride Priamo Elena.
User: Paride.
System: In Greek mythology, Paride was the child of Priamo and Ecuba; just before his birth, his mother dreamed that she would give birth to a flaming torch …..
User: Skip.
System: Are you interested in another character? …..
User: No.
System: Great, let's move to the next scene.

Figure 15. Dialogue in the scene perspective

This approach provides user with a virtual guide that assist the virtual tour according to him/her preferences. Work is in progress to enable voice navigation of the virtual environment. Using his/her voice, the user would issue commands for simulated flight or "teleportation" to arbitrary locations without the need to navigate step-by-step , find entities in a large 3D space, etc.

7. CONCLUSIONS

This paper proposed the aim of experimentation of innovative methodologies for the survey and fruition of Cultural Heritage.
With reference to the problems of the survey, particularly, integration among different types of laser-scanning instruments and between laser-scanning and photogrammetric techniques has been studied.
In regard to the aspect of the fruition, which today has become of greater importance with respect to the past, a system was tested directed to the visualization and navigation by WebGIS techniques of 3D rendered model previously produced, to allow a more widespread accessibility to the cultural heritage. With the use of this system it is possible to generate, navigate and explore reconstructed environments of cultural interest, enabling the extended fruition of works of art which are not physically accessible to the user. The system embed the virtual environment with a multimodal interactive guide which assists the user during his virtual tour. Visitors can navigate the virtual representation of the ancient wood ceiling with tempera paintings and achieve, interacting vocally, relevant meta information about history and background of painted scenes.

REFERENCES

References from Books:
Bologna F., 2002. *Il soffitto della Sala Magna allo Steri di Palermo*, Flaccovio, Palermo (Italy).

Building a Better Golem, Markup Languages: Theory and Practice 2.4, September 2000, (pp. 337-351) Copyright 2001 by the Massachusetts Institute of Technology.

References from Other Literature:
Agnello F., Corsale R., Franco V., Lo Brutto M., Midulla P., Orlando P., Villa B., 2003. Cultural heritage and information systems, an investigation into a dedicated hypertext. *Vision techniques for digital architectural and archaeological archives*, Ancona, 1-3 luglio 2003, Volume XXXIV, part 5/W12, pp. 7-12.

Alshawabkeh, Y., Haala, N., 2005. Automatic multi-image photo-texturing of complex 3D scenes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, Volume XXXVI-5/C34, pp. 68-73.

Beraldin, J.A., Picard, M., El Hakim, S.F., Godin, G., Latouche, C., Valzano, V., Bandiera, A., 2002. Exploring a Byzantine crypt through a high-resolution texture mapped 3D model: combining range data and photogrammetry. *Proceedings of the CIPA WG6 International Workshop on Scanning for Cultural Heritage Recording*, Corfù, Greece, 1-2 September 2002, pp. 65-72.

Böhm, J., 2004. Multi-image fusion for occlusion-free façade texturing. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, Volume XXXV-5, pp. 867-872.

D'Amelio S., Emmolo D., Lo Brutto M., Orlando P., Villa B.,2005. 3D Techniques for the survey of cultural heritage. *Proceedings of XXth CIPA International Symposium,* pp. 215-219.

Don Brutzman, "X3D-edit authoring for extensible 3D (X3D) graphics", *ACM SIGGRAPH 2003 Educators program from the*

30th annual conference on Computer graphics and interactive techniques, July 27-31, 2003, San Diego, California.

Gonzo L., Voltolini F., Girardi S., Rizzi A., Remondino F., El-Hakim S.(2007). Multiple Techniques Approach to the 3D Virtual Reconstruction of Cultural Heritage. Invited Paper. *5th Eurographics Italian Chapter Conference 2007* - R. De Amicis and G. Conti (Editors), 14-16 Feb 2007, Trento, Italy.

Haiqing Li, Guofu Yin, Jie Fu, "Research on the Collaborative Virtual Products Development Based on Web and X3D", *Proceeding of the 16th International Conference on Artificial Reality and Telexistence* - Workshops ICAT, 2006. Nov. 2006 Page(s):141 – 144.

Salonia P., Bellucci V., Scolastico S., Marcolongo A., Messina T.L. (2007). 3D Survey Technologies for Reconstruction, Analysis and Diagnosis in the Conservation Process of Cultural Heritage. *XXI International CIPA Symposium*, 01-06 October, Athens, Greece.

Petsa E., Grammatikopoulos L., Kalisperakis I., Karras G., Pagounis V. (2007). Laser Scanning and Automatic Multi-Image Texturing of Surface Projections. *XXI International CIPA Symposium*, 01-06 October, Athens, Greece.

Scianna A, Villa B., Ammoscato A, Corsale R., 2004. Multimedia guide in archaeological sites by GIS-GPS techniques, pocket PCs and pocket GPSs. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* Volume XXXV, B2, pp. 635-640.

Trumello E., Santangelo A., Gentile A., Gaglio S. (2008). "A Multimodal Guide for Virtual 3D Models of Cultural Heritage Artifacts." *Proceeding of the first International Workshop on Intelligent Interfaces for Human-Computer Interaction* (IIHCI), Barcellona, March 4-7, pp. 735-740.

Valzano V., Bandiera A., Beraldin J.-A., Picard M., El-Hakim S., Godin G., Paquet E., Rioux M. (2005). Fusion of 3D Information for Efficient Modeling of Cultural Heritage Sites with Objects. In *International Archives of the Photogrammetry, Remore Sensing and Spatial Information Sciences*, vol. XXXVI-5/C34-CIPA.

References from websites:
"X3D Specifications, Encodings and Language Bindings", Web3D Consortium, Inc. Available at: http://www.web3d.org/x3d/specifications/

Acknowledgements
We wish to thank:
- CAM2 s.r.l. for making the FARO LS880 scanner laser available;
- Ing. Gianni Giordano for supporting surveying operations and the processing of the digital model;
- Architects Giacomo Carpino, Carmelo Clementi, Alessandro Corso, Liria Di Maria, Marcella Mannina and Viviana Trapani, students of the Master in "Surveying and representation of cultural heritage" (2005-06), for having actively participated to the surveying and modelling processes.

3D MAPPING OF CULTURAL AND NATURAL HERITAGE: THE CASE STUDY OF THE CAVE OF POLYPHEMUS CYCLOPE

Patias, P., Sylaiou, S., Georgiadis Ch., Georgoula, O., Kaimaris, D., Stylianidis, S.

Laboratory of Photogrammetry and Remote Sensing, Faculty of Rural & Surveying Engineering, The Aristotle University of Thessaloniki, Univ.Box. 473, 54124, Greece, patias@auth.gr,
Tel +30-2310-996116, Fax +30-2310-996128, Univ Box 473, GR-54124 Thessaloniki

KEY WORDS: cave, recording, laser scanning, visualisation, cultural informatics, e-Heritage

ABSTRACT:
This study describes the creation of a three dimensional model for the cave of Polyfimos in the prefecture of Rodopi, Greece. The cave is of historical significance, because according to archaeologists there are traces of its usage going back to the prehistoric era. Its name was given from the myth concerning the blinding of Cyclops Polyfimos, a one-eyed giant in Homer's Odyssey. Currently there is a question for the touristic exploitation of the cave. The atmospheric conditions inside the cave must be carefully studied, in order to preserve the fragile biological conditions, the fauna and flora. In order to achieve the detection of such openings of the natural ventilation system, a three dimensional mapping of the area with the aid of laser scanner for the three dimensional modeling of the cave's interior and aerial photogrammetry, as well as differential GPS measurements for the mapping of the surrounding area are essential.. The research challenge is the fusion of the data using these different measurement methods for such a complex object.

1. INTRODUCTION

1.1 Monuments in danger

Despite the establishment of international organisations, as UNESCO (*United Nations Educational, Scientific and Cultural Organization)*, ICOM (*International Council of Museums*), ICBS (*International Committee of Blue Shield*), ICOMOS (International Committee of Monuments and Sites) etc. and the signature of relevant protocols, as the *Geneva Convention of 1949 and its two protocols, the Hague Convention for the Protection of Cultural Property on the Event of Armed Conflict του 1954, the Venice Charter of 1964 for the Conservation and Restoration of Monuments and Sites, the UNESCO Convention of 1970 on the Means of Prohibiting and Preventing the Illicit Import, Export and Transfer of Ownership of Cultural Property, Convention concerning the Protection of the World Cultural and Natural Heritage* etc., natural and cultural monuments are vulnerable from a number of dangers that we could classify them to the following basic categories:

- *Physical destructions*, as earthquakes, extreme weather phenomena and Time.
- *Destructions caused by accidents,* as fires.
- *Man-made threats*: political conflict, looting, unchecked urban and industrial development, pollution, tourism, and, for the first time, global climate change (Evarts, 2007).
- *The destructive nature of archaeological excavation* that imposes the removal of the upper units of stratification.

Monuments of Natural and Cultural Heritage, such as the caves, especially when they have a long history and are closely connected with human presence and activity over the centuries need special attention and must be preserved and carefully recorded.

1.2 The motivation of the study

In our case the cave of Cyclops Polyphemus, mentioned by Homeric Odyssey, which is situated in Maronia (Thrace, Greece) was recorded with the aid of a ground-based laser scanner in order to produce an accurate and realistic 3D model of the cave in an efficient and cost-effective way. Furthermore, the Digital Terrain Model of the area was produced and connected with the 3D model of the cave.

The cave has a long history with traces that detect the human presence from the prehistoric years and more specifically from the Neolithic period (Panti and Myteletsis 2006). Also, it has great interest and importance from a geological point of view (Lazaridis, 2005) and it is the habitat of various animal species (Tsompanikolidis, 1998).

The last years there is a pressure for tourist development of the cave. This can be highly problematic, because the atmospheric conditions can change and provoke problems to the equilibrium of the cave.

1.3 Aim of study

As it is already mentioned the conditions of the cave are very fragile and they must be taken into account and carefully protected when the cave will be open to the wide public. Most of the times, the main problems of the caves are a consequence of anthropogenic impact. More specifically, researches have shown that the massive influx of visitors can provoke changes to the cave microclimatology, like temperature and pressure gradient and humidity, by the emission of carbon dioxide and water vapour (Villar et al., 1984; Fernández et al., 1986; Bower, 1986; Pulido-Bosch et al. 1997; Sanchez-Moral, 1999; Liñán et al., 2008). In the case of the cave of Polyphemus, these changes can put the cave in risk and cause dissolution of the rocks and disturbance of the sheltered bats. This realisation of this scenario will cause the departure of the animals and the transformation of the cave to another dead cave with no animals (NATURA 2000).

The microclimate and the atmospheric conditions of the cave need to be preserved, as well as the animals of the cave need to be protected. In order for the cave to be open to public, future problems in ventilation deterioration caused by the visitors must be resolved. By ventilating the room using the natural ventilation openings, the rooms can return to their prevailing conditions before visitors were allowed to get in.

The best way to control the ventilation is to find the natural openings and utilize them, in order to make measurements in temperature, humidity and CO_2 emissions, so as to calculate the visitor capacity per hour or day preserving thus the current condition of the site where it is possible and not permitting the visit to rooms where this is not feasible. Furthermore, according to other studies, these measurements can provide a valuable help in calculating also the amount of time that the groups of visitors can spend in each room and in finding the best period of time for visits in the cave, in which the greatest ventilation is produced and at that time the cave has the greatest capacity for recovery (Hoyos et al. 1998).

The natural holes, which allow air circulation inside the cave, shall be located, in order to aid the future researchers making the appropriate air measurements to preserve the cave's microclimate. In this paper we propose a virtual presentation of the cave rooms that can be visitable via Internet or a 3D virtual reality theatre. The main aim of our study was to create the three dimensional model of the cave's interior and exterior. In addition, a series of sections and cross sections were produced, giving more details for certain areas inside the cave. In this paper the methodology undertaken for the 3D mapping of the cave and the final result are described, the related problems are presented and proposals for future work are investigated.

2. LITERATURE REVIEW

2.1 Laser scanning in caves

Researches that were undertaken showed that there are not a lot of applications of caves visualisation with the aid of laser scanners. Between these cases (IKEUCHI Laboratory) the most known is that of the Altamira cave in northern Spain (Donelan, 2002) where a scanner Minolta VI-700 was used for the protection against further deterioration of the cave paintings caused by the mass coming of visitors. From 1879 till 1978 almost 177.000 people visited the cave and provoked major changes to the atmospheric conditions and consequently serious decays to the paintings. The cave was recorded with the aid of a laser scanner and its 3D model was produced. A series of procedures, in which scientists with various backgrounds, as well as sculptors and painters were collaborated, in order to produce the exact negative shapes of the cave that were cut in foam and then it was made and exposed in a museum space in Spain a replication of the cave in physical size from silicone moulds (Minolta 3D Applications in Restoration, Conservation and Cultural Heritage, RapidForm).

Another application that used laser scanner, in conjunction with Virtual Reality technologies is in the caves Dunhuang in China (Lutz and Weintke, 1999). According to designs and onsite measurements a 3D reconstruction of the cave was created and details that were produced by photogrammetric and laser scanner methods were added. The 3D model was used not only for the documentation and the digital preservation of the cave, but also for its presentation to the wide public. The research products were used for the creation of a Virtual Environment CAVE®, a room that its walls are used as projection screens. In this environment the visitor can wear special glasses that permit immersion, interaction and navigation to the virtual cave (Lutz and Weintke, 1999).

In another project the Grotta dei Cervi cave, a complex and fragile Neolithic cave with pictographs and petroglyphs was recorded and its details were well documented (Beraldin et al. 2006). As future work it is intented to use the project results for monitoring the condition and stability of the fragile cave. Furthermore, it is proposed to present the detailed 3D model to a 3D VR Theatre, in order to enhance the understanding of the site and to raise awareness (ibid).

3. THE CYCLOPS POLYPHEMUS CAVE

3.1 Geographic location

The cave can be found at a distance of 30 klm. south of the Komotini city in the prefecture of Rodopi, between the villages Proskynites and Maronia, near the mountain Ismaros and south of the river Platanitis (Figure 1) (40° 56' 40'' N, 25° 35' 0'' E).

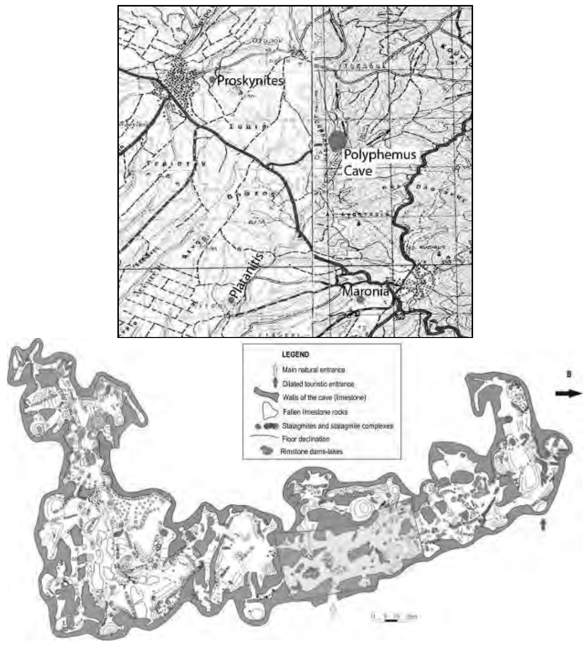


Figure 1: Map of the greater area and the cave

The cave is located in a limestone hill and its access is through a 2 km dirt road. The cave's entrance is not accessible by a vehicle. The distance of the entrance from the road is approximately 250 m. In order to reach the entrance one must walk through a trail and descend a relative steep slope (Figure 2). The inside temperature is constant at about 16^o C.

Figure 2: Entrance of the cave as seen from the hill

The cave's entrance has a height of 1.70 m leading to a steep descending trail that ends in the front chambers of the cave, those of the Altar and the Throne. The two chambers are separated by a big pillar (Figure 3).

Figure 3: Pillar that separates the two chambers

3.2 Physical features

The cave is protected by *NATURA 2000 Network* (GR1130008). It is a habitat for 11 protected species of bats and a shelter for 31 species of invertebrates (Tsombanikolidis, 1998; Beron et al., 2004; Melfos et al., 2005), because of the ideal conditions of the inner environment of the cave and the existence of running water close to it and of two small lakes inside it.

3.3 Geological features

Polychromatic stalactites and stalagmites are in abundance inside the cave (NATURA 2000). A study concerning the stability of the cave was undertaken and its main purpose was:

- to determine any type of potential failure and
- to propose the more appropriate supporting methods taking into account the already existing not only the safety of visitors, but also the physical features of the cave (Christaras et al., 2004).

The tectonic data of the cave recorded and tectonic and stability diagrams were created, in order to determine the stability factors regarding to wedge and planar failures. The researchers proposed an arrangement of stainless pre-tensioned self-drilling rockbolts for the reinforcement support of the roof (ibid).

3.4 Archaeological interest

The cave of Polyphemus (Thrace, Greece) was connected by tradition with the cave of Cyclops Polyphemus (Greek Κύκλωψ Πολύφημος), the son of Poseidon and a sea nymph Thoosa, a giant with a single eye in the middle of its forehead. It was in this cave that he get drunk and then was poked his eye out with a stake by Odysseus, as mentioned by Homeric Odyssey (ι352-452).

In 1938 the archaeologist Yeoryios Bakalakis identified the cave with the cave in which Orpheus lived in the region of Cicones, according to a passage in Apollonius' Argonautica (Liapis, 1993). The cave has been excavated by the Ephorate of Palaeoanthropology- Speleology Northern Greece for the last two years. The archaeological remains reveal human presence in its interior from Neolithic since Late Byzantine period (Panti and Myteletsis, 2006; Panti and Bachlas, 2007).

4. STEPS UNDERTAKEN FOR THE 3D MAPPING

4.1 Data

The cave is approximately 350 m long and has a width between 15 and 50m. Due to limitation of time and economic constraints it was not feasible to scan all the rooms of the cave. The part of the cave that was scanned was the chamber of the Altar and the chamber of the Throne.

4.2 Instrumentation

4.2.1 *Laser Scanner Trimble GS200*

A 3D Laser Scanner Trimble GS200 was used for the recording of the cave. It offers a field of view of 360° horizontal and 60° vertical. Its sensor permits measurements of an accuracy of 1.5mm in a 50 m distance with a laser beam of 3mm diameter. Besides the recording of the coordinates X, Y, Z of the points, there is the possibility of recording colour information (R, G, B), as well as of the returned intensity.

4.2.2 *Total station Leica TPS400*

A total station TPS403 was used for the measurement of the όδευσης and the points inside the cave. The mini reflector with circular prism GPR1 with a measuring range of 3.5 km and an accuracy of 2 mm + 2 ppm was used. The measurement of points inside the cave was done reflectorless with an accuracy of 3 mm + 2 ppm for the baseline measurements and a range of 170m.

4.2.3 *Additional instruments*

One laptop compatible with the minimum demands for the control of the Laser Scanner was used. A GPS unit -Thales Promark3- was used in combination with a permanent GPS Station EUREF (AUT1) located at the Aristotle University of Thessaloniki and used as a reference. The solution of the two (2) GPS points was accomplished by using GNSS Solutions software.

Furthermore, two (2) head lamps were used for the lighting of the cave, as well as an electric generator and extension cords for transmitting electricity inside the cave.

5. METHODS AND RESULTS

5.1 Measurement procedure

Four (4) days of measurement on-site needed for the documentation of the Polyphemus cave. A Laser Scanner Trimble GS200 was used for the collection of the point cloud. For the recording of the external positions and the georeferencing of all data in the Greek Geodetic Reference System 1987 (ΕΓΣΑ 87, GGRS 87) a Geodetic Station was used, a system for measuring angles and distances of reference points on the cave, which are further transformed to coordinates for the measurements that connect the internal and the external space of the cave and GPS receivers.

Inside the cave seven (7) stations were used by the Laser Scanner for the collection of the point cloud. The selection of the station was made taking into consideration the limitation of the vertical field of view, the 60o of the system, in order not to produce "gaps" to the final point cloud and also to ensure the desirable resolution of the 15mm between the points. The measurements were made with high accuracy (High Quality in Pointscape) and for the acquirement of the color information (RGB) of every point the automatic self focusing of the Laser Scanner was used increasing the time of reception and staying to the field. For the registration of the seven (7) point clouds seven spheres were used as targets. Every two point clouds there were at least three (3) common target points and by applying the point clouds registration to the software HDS Leica Cyclone with a final model accuracy of 1 cm.

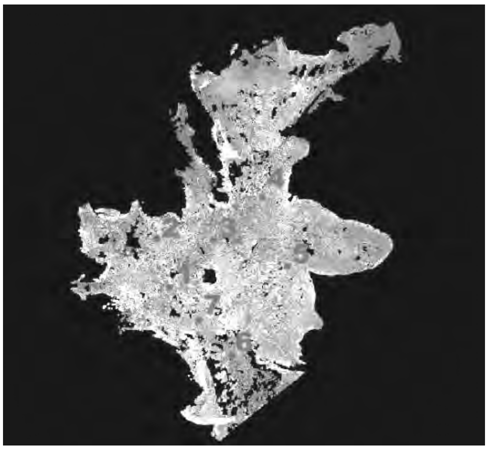


Figure 4: Geodetic transverse stations

Four stations were measured by the Geodetic Station for the georeferencing of the final point cloud from the cave, two of them to the surrounding area of the cave (S1 andS2), one at the entrance (S3) and one in the internal area of the cave (S4) (Figure 4). From the internal position (S4) thirteen flat targets were recorded inside the cave. The same targets were also recorded from the Laser Scanner during the collecting of the point cloud. Initially, the network was solved and the three dimensional coordinates of positions and the targets were calculated in an independent coordinate system. Then, after the solution of the position measurements with the aid of GPS, they were transformed to the Greek Geodetic Reference System 1987 (ΕΓΣΑ 87, GGRS 87).

The DTM was collected using the topographic network that was established in the area of the cave. The measurements were realized using a Leica Total Station. For the demands of our study points on grid of approximately 1x1m were measured in a 50x50 m area. The coordinates of the points were computed in GGRS 87.

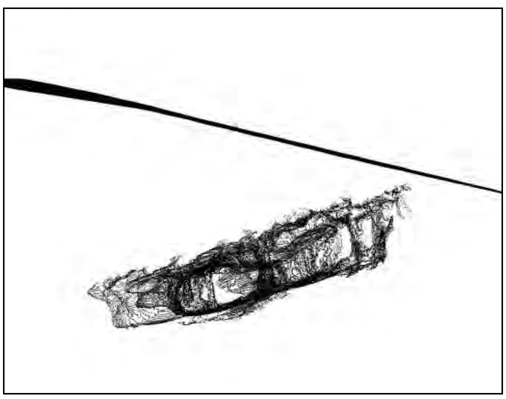


Figure 5: Cave cross section and the natural earth profile

For the georeference of the point cloud ten (10) points were used. These points were measured using a total station and their coordinates were transformed from the local coordinate system to the GGRS coordinate system. The georeferencing procedure was performed using the Polyworks software. Then, the georefenced point cloud and the DTM were united in Polyworks and a series of cross sections were created in order to study the cave. In Figures 5 and 6 we can see two different cross sections along the X and Y axes.

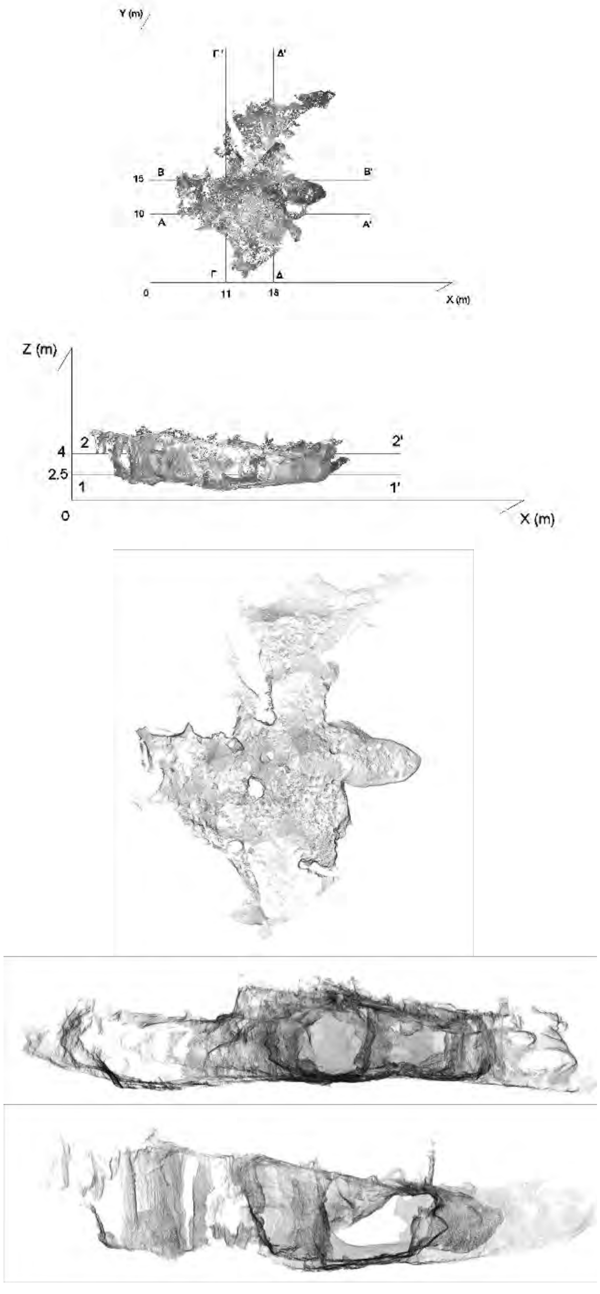


Figure 6: Indicative horizontal and vertical cross sections

6. CONCLUSIONS AND FUTURE WORK

This paper presents a summary of the 3D mapping work that was accomplished in order to create the 3D model of the cave area. Despite some difficulties in recording the irregular surface shape of the rock surface and speleothems , the 3D model of the cave can provide future solutions in order to make the parts of the cave that must be hidden from view with fragile biological equilibrium and protected species of animals, available to the public. An immersive 3D virtual reality theatre located near the cave, or a website with the embedded 3D model of the cave would enable visitors to “virtually” visit the site (Beraldin et al., 2006). Researchers can magnify or zoom in on a 3D model to examine measure and compare fine surface details for signs of deterioration or to examine tool mark features.

7. REFERENCES

Beraldin, J. A., Blais, F., Cournoyer, L., Picard, M., Gamache, D., Valzano, V., Bandiera, A., and Gorgoglione, M., 2006. Multi-Resolution Digital 3D Imaging System Applied to the Recording of Grotto Sites: The Case of the Grotta dei Cervi. In: *Proceedings of the 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2006)*, http://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-48791.pdf (accessed: 10 June 2008)

Beron P., Petrov B., and Stoev P., 2004. The Invertebrate Cave Fauna of the Eastern Rhodopes (Bulgaria and Greece). In: Biodiversity of Bulgaria. 2. Biodiversity of Eastern Rhodopes (Bulgaria and Greece). Pensoft & Nat. Mus. Natur. Hist., Sofia, 791-822.
http://hinko.org/hinko/Dowloads/Eastern%20Rhodopes.pdf (accessed at: 10 June 2008)

Bower B., 2986. Altamira cave art: low-vent district? - cave painting damage caused by carbon dioxide exhaled by tourists. Science News. May 31, 1986. FindArticles.com. http://findarticles.com/p/articles/mi_m1200/is_v129/ai_427749

Casale, A., 1984. Una nuova Muelleriella delle Grecia settentrionale (Coleoptera, Catopidae, Bathysciinae). *Doriana* 6(253): 1-6.

Christaras, B., Chatzigogos, Th., Dimitriou, A., Chatzigogos, N., Makedon, Th., Vogiatzis, D., and Filippidis, S., 2004. Stability of Polyfimos Cave in Maroneia, Thrace. In: *Proceedings of the 10th International Congress*, Thessaloniki, Greece. http://www.geo.auth.gr/ege2004/articles/TG37_374.pdf (accessed at: 10 June 2008)

Donelan, J., 2002. Making Prehistory. *Computer Graphics World*, March 2002, pp. 32-33.

Evarts H., 2007. 2008 World Monuments Watch list of 100 most endangered sites.
http://www.wmf.org/pdf/Watch_2008_release.pdf (accessed at: 10 June 2008)

Faimon, J., Štelcl, J., Zimák, J., Zajíček, P., and Schwarzová, M., 2006. Speleothem destruction in the caves of the Moravian Karst (Czech Republic).
http://www.ssj.sk/pdf/VVOJ_05/Z05_25.pdf (accessed: 10 June 2008)

Fernández, P. L., Gutierrez, I., Quindós, L. S., Soto, J., and Villar, E., 1986. Natural ventilation of the Paintings Room in the Altamira cave. *Nature* 321, pp.586-588

Hoyos, M., Soler, V., Cañaveras, J., Sánchez-Moral C. S., and Sanz-Rubio, E., 1998. Microclimatic characterization of a karstic cave: human impact on microenvironmental parameters of a prehistoric rock art cave (Candamo Cave, northern Spain). *Environmental Geology* 33(4), March, 1998.

IKEUCHI Laboratory, Computer Vision Laboratory, Institute of Industrial Science, University of Tokyo, Fugoppe Cave Site. http://www.cvl.iis.u-tokyo.ac.jp/gallery_e/Fugoppe/fugoppe.html (accessed at: 10 June 2008)

Koutsaftikis, A., 1974. Recent butterfly records from Greece. *Entomologist's Record and Journal of Variation*, 86, pp. 15-17

Lazaridis, G., 2005. Observations about the shields of the cave of the Cyclops Polyphemus in Maronia (Prefecture of Rodopi) (in greek). *Bulletin of the Geological Society of Greece* 37. http://geolib.geo.auth.gr/ojs-2.1.1/index.php/bgsg/article/view/1218/1088 (accessed at: 10 June 2008)

Liapis, A., 1993. Komotini and the surrounding area. The major archaeological sites,
http://alex.eled.duth.gr/komotini/index.html

Liñán, C., Vadillo, I., and Carrasco, F., 2008. Carbon dioxide concentration in air within the Nerja Cave (Malaga, Andalusia, Spain). *International Journal of Speleology*, 37 (2), pp. 99-106.

Lutz, B., and Weintke, M., 1999. Cultural Heritage: Virtual Dunhuang Art Cave: A Cave within a CAVE. *Computer Graphics Forum*, 18(3) (September 1999).

Melfos, B., Chadjipetros, A., Chatzopoulou, A. Vasiliadou, A., Lazaridis G., Vaxevanopoulos, M. Syridis, G., Tsoukala, E., and Pavlidis S., 2005. Geological, petrological, palaeontological study at the Maronia cave in the eocene, nummulitic limestone in western Thrace (in greek). *Bulletin of the Geological Society of Greece* 37.
http://geolib.geo.auth.gr/ojs-2.1.1/index.php/bgsg/article/view/1220/1090 (accessed at: 10 June 2008)

Minolta 3D Applications in Restoration, Conservation and Cultural Heritage,
http://www.minolta3d.com/applications/pdfs/culturalheritage.pdf (accessed at: 10 June 2008)

NATURA 2000, Maronia Cave (GR1130008), Network of researchers in environmental management, Hellenic Ministry of Environment, physical planning and public works, Greece, http://natura.minenv.gr/natura/server/user/biotopos_info.asp?siteCode=GR1130008&lng=GR

Panti, A., and Myteletsis, M., 2006. Excavation to the cave “Polyphemus” of Maronia, S. Rodopi: Preliminary research. In: *Proceedings of the Archaeological Work accomplished in Macedonia and Thrace (AWMT)* 20, (in greek, under publication).

Panti, A., and Bachlas A., 2007. The excavational research in cave Polyfimos during the year 2007. In: *Proceedings of the*

Archaeological Work accomplished in Macedonia and Thrace 20, 2006 (AWMT) (in greek, under publication).

Pentazos, Ev., 1973. Statement report from Athens Archaeological Association in 1971. Athens.

Pulido-Bosch, A., Martin-Rosales, W., Lopez-Chicano, M., Rodriguez-Navarro, C.M., and Vallejos, A., 1997. Human impact in a tourist karstic cave (Aracena, Spain). *Environmental Geology*, 31 (3/4), pp. 142-149.

RapidForm, 3D Scanning Solutions, Cultural Heritage, http://www.rapidform.com/Contents/Solutions/category_id/39 (accessed at: 10 June 2008)

Sanchez-Moral, S., Soler, V., Canaveras, J. C., Sanz-Rubio, E., Van Grieken, R. and Gysels, K., 1999. Inorganic deterioration affecting the Altamira Cave, N Spain: quantitative approach to wall-corrosion (solutional etching) processes induced by visitors. *The Science of the Total Environment*, 243, pp. 67–84.

Tsompanikolidis, K., 1996 (thesis). Personal observations on systematic and population study on chiroptera of cave Maronia, Rodopi, Northern Greece, Department of Biology, The Aristotle University of Thessaloniki, Greece.

Villar E, Bonet A, Díaz B, Fernández PL, Gutiérrez I, Quindós LS, Solana JR, Soto J (1984) Ambient temperature variations in the Hall of Paintings of Altamira Cave due to the presence of visitors. *Cave Science* 11, pp. 99–104

8. ACKNOWLEDGEMENTS

The authors acknowledge the financial support for this pilot project was provided by the Greek Ministry of Culture and the Department of Surveying Engineering of the Aristotle University of Thessaloniki for providing the laser scanning equipment. The authors would also like to express their gratitude to Miss E. Poulaki, Director of the Ephorate of Antiquities of Pieria, ex-Director of the Ephorate of Palaeoanthropology and Speleology for her continuous interest and support and Dr. E. Veleni, Directress of the Archaeological Museum of Thessaloniki, Greece, also ex-Director of the Ephorate of Palaeoanthropology and Speleology for making this collaboration possible. Great thanks to Thanasis Georgantas, Aggelos Tsakiroglou and Christos Stamatopoulos, surveying engineers that operated the laser scanner, processed the data and created the first 3D model of the cave. We wish to thank Anna Panti and Miltiadis Myteletsis, archaeologists in guiding the authors to the archaeological site and provided their assistance during the data acquisition. Finally, we would like to acknowledge the ex-mayor of Municipality of Maronia, Mr Topouzli, who generously provided technical support during the cave survey.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

RECONSTRUCTING 3D FACES IN CULTURAL HERITAGE APPLICATIONS

Andreas Lanitis[a] and Georgios Stylianou[b]

[a] Dept. of Multimedia and Graphic Arts, Cyprus University of Technology , P.O. Box 50329, 3036, Lemesos, Cyprus,
andreas.lanitis@cut.ac.cy
[b] Dept. of Computer Science, European University Cyprus , P.O. Box 22006, 1516, Nicosia, Cyprus,
g.stylianou@euc.ac.cy

KEY WORDS: Image-Based 3D Face Reconstruction, 3D Face Shape Models

ABSTRACT:

Image-based 3D face reconstruction techniques can be used for generating a 3D model of a person's face given a single or multiple images of the face of a person. Such techniques could be used for generating 3D models of faces appearing in cultural heritance artefacts, based on images captured with ordinary digital cameras. The ability to generate 3D reconstructions of faces appearing in cultural heritage artefacts will facilitate the development of numerous applications such as the generation of cultural heritage related 3D animations, short films, educational applications and customized educational games. In this paper we review key image-based 3D face reconstruction algorithms belonging to different categories and assess their applicability for reconstructing 3D faces appearing in cultural heritage artefacts. We also describe a dedicated 3D face reconstruction algorithm that can be used for efficient reconstruction of faces in cultural heritage applications based on a single image. The proposed algorithm requires the location of a number of landmarks on the face to be reconstructed. A deformable 3D shape model is used for generating an approximation of the 3D face geometry of the face shown on the artefact and texture mapping is used for generating the full 3D model. In the paper we present initial quantitative results and visual results obtained when the method was used for reconstructing 3D faces appearing in different types of cultural heritage artefacts such as statues, paintings, murals and mosaics.

1. INTRODUCTION

Ancient sites and artefacts such as wall paintings, paintings, statues, pots, anaglyphs, mosaics and icons contain a plethora of human faces that usually display important persons in history (see figure 1).

Figure 1: Faces appearing in different types of CH artefacts

In this paper we review methods that can be used for generating three-dimensional (3D) reconstructions of faces appearing in such artefacts. The ability to generate a complete 3D models of faces appearing in cultural heritance artefacts can be very important in numerous applications such as:

- 3D animations: Once 3D facial models are created, it is trivial to use animation techniques for generating animations that can be used in short films and other educational applications.
- Educational Games. 3D reconstructed faces can be used for generating realistic characters appearing in CH related educational games.
- Art related studies: The generation of the actual 3D face geometry of subjects appearing in CH artefacts will allow the comparison of faces of the same subject generated by different artists, so that comparison of artistic methodologies and trends will be enabled. The definition of such trends will be most important for studying the design methodologies and styles adopted in different countries and different time periods.
- Digital restoration: 3D reconstructions of damaged faces shown in artefacts can be used as the basis for digital restoration of the facial appearance of missing/damaged parts.

Although it is possible to generate a 3D model of a face using dedicated 3D scanners, this method is only applicable in applications involving the digitisation of existing 3D facial structures i.e. faces of statues or faces in anaglyphs. 3D scanners cannot be used for generating 3D models of faces appearing as 2D surfaces such as faces appearing in paintings, mosaics and murals. Even when dealing with the generation of 3D models of statues, the use of 3D scanners possesses several limitations due to the size and cost of 3D scanners, the requirement of controlled lighting during the scanning process, the considerable amount of post processing required for removing noise and joining 3D scans, and the overall processing time required during the scanning process. Such problems are minimized when on-site data acquisition is performed using 2D instead of 3D imaging equipment.

As an alternative to using 3D scanners, image-based 3D reconstruction techniques can be used for generating 3D face models of faces appearing in CH artefacts. Over the past decade the technology of image-based 3D face reconstruction has evolved dramatically. Realistically looking 3D faces can be reconstructed from a single or more photographs including video frames (Stylianou, 2008). However, most 3D face reconstruction techniques reported in the literature aim towards the reconstruction of real faces appearing in digital face images. To the best of our knowledge as of today no systematic efforts in reconstructing 3D faces appearing in CH artefacts were reported in the relevant literature.

In this paper we provide a general overview of 3D face reconstruction methods reported in the literature. In particular we describe typical example-based (Blanz, 1999), stereo-based (Akimoto, 1993; Chen 2001) and video-based (Leung, 2000) 3D reconstruction methods. For each category of 3D reconstruction algorithms we provide a discussion that assesses the applicability of typical algorithms for reconstructing faces appearing in CH artefacts. We also propose a dedicated method for 3D face reconstruction of CH artefacts and preliminary results when the method is applied for reconstructing faces appearing in different types of artefacts.

2. 3D FACE RECONSTRUCTION METHODOLOGIES

Image-based 3D face reconstruction methods reported in the literature can be classified as example-based, stereo and video-based methods. Usually example-based methods operate on a single input image, stereo methods operate on a pair of images and video-based methods operate on multiple input frames. In this section we provide a brief overview of typical 3D face reconstruction approaches for each category, outline the major advantages and disadvantages of each category and discuss the applicability of those approaches for reconstructing faces appearing in CH artefacts. The review provided in this section aims to present the basic principles pertaining to 3D reconstruction methods - a more comprehensive review appears elsewhere (Stylianou, 2008).

2.1 Example based Methods

Example-based methods rely on the analysis of a large number of training samples for accomplishing 3D reconstruction so that the structure and typical deformations associated with 3D faces are learned. The information obtained during the training process is used to guide the reconstruction process.

An example-based 3D face reconstruction method has two phases: The morphable model generation phase and the actual reconstruction phase. During the morphable model generation phase the training set, which is composed of a face database containing 3D face models, is used for building a parametric face model, capable of generating 3D faces consistent with samples appearing in the training set (Atick, 1996; Blanz, 1999). Once a parametric face model is trained, it is possible to generate different 3D faces by setting values to the model parameters. 3D face reconstruction is accomplished by defining optimum values for the face model parameters so that a 3D face generated by the model resembles the face appearing in a given image. The actual steps involved in this process are:

1. The values of the model parameters are initialized to zero so that the model generates the mean 3D face instance.
2. Based on the values of the model parameters the corresponding 3D face instance is generated.
3. The 3D face instance is projected on the face in a given image.
4. The similarity between the given face and the 3D model instance is estimated. Usually the similarity measure is defined based on differences in texture between the 3D model instance and the face to be reconstructed.
5. An optimization algorithm is used for defining improved values for the model parameters.
6. Steps 2 to 5 are repeated until the similarity between the 3D face instance and the given face is maximized.

When the similarity measure is maximised a 3D face with appearance similar to the face in the given image is created so that the task of reconstructing a 3D face from a single face image is accomplished.

The most cited example-based 3D reconstruction methodology is the method developed by Vetter and Blanz (Blanz, 1999) where a PCA based model is used. In this case 3D reconstruction is achieved by computing the weights of the shape and texture eigenvectors for a novel 2D face image. The method requires manual initialization to roughly align the 3D morphable model and the novel face image. During the reconstruction process 22 rendering parameters such as pose, focal length of the camera, light intensity, color and direction are estimated. Other typical example-based methods are reported by Hu et al (Hu, 2004), Lee et al (Lee, 2005) and Romdhani (Romdhani, 2005).

The main advantage of example-based 3D reconstruction methods is the ability to produce visually good 3D face reconstructions based on a single face image. The use of a face model allows the implicit use of constraints related to the expected structure of a human face, so that it is possible to generate full 3D reconstructions based on missing data (i.e. data seen only from a single face view). However, example-based techniques exhibit the following disadvantages:

- Limited anatomical accuracy of generated models. This is reasonable as information from a single image is not enough for obtaining a highly accurate estimate of the 3D face structure.
- Due to the high computation requirements involved, example-based techniques usually do not run in real time.

In the majority of example-based methods reported in the literature, the calculation of the similarity between a 2D projected face model instance and the face to be reconstructed mainly relies on texture differences between corresponding pixels. For this process to be successful it is assumed that faces in the given image have skin-like texture similar to facial texture encountered in digital images. For this reason this method is not applicable to reconstructing faces appearing in CH artefacts since in such cases irregular facial textures are encountered according to the type and material used for designing the artefacts (see figure 1).

2.2 Stereo Methods

Stereo 3D reconstruction methods can be classified to two main categories: (a) methods based on orthogonal images and (b) methods based on non-orthogonal images. In category (a) it is necessary to have two images each containing a face in frontal and profile views, respectively. In the second category, it is necessary to have at least two calibrated cameras (between them) in order to capture two images of the same face from different viewpoints. Both methods identify corresponding features from the input images but use them differently as described in sections 2.2.1 and 2.2.2 respectively.

2.2.1 Orthogonal images. Reconstruction algorithms based on orthogonal images rely on the location of common features on both frontal and side photos that correspond to predefined features on a generic head model. The location of common features enables the calculation of the x and y coordinates of each vertex using the frontal image and the calculation of the z coordinate of each vertex based on the profile image. Once the coordinates of a number of vertices are defined it is possible to deform a generic head model in order to approximate the geometry defined by the extracted vertices. Once the geometry of a 3D model is defined, the face texture is generated from both images and texture mapped onto the resulting 3D model. Typical orthogonal-based 3D reconstruction methods are reported by Akimoto (Akimoto, 1993), Lee (Lee, 1997) and Park (Park, 2005).

Ortogonal-based methods can produce geometrically accurate faces since information from orthogonal images provides an adequate description of the 3D geometry of a face. Even though only two photographs are used it is trivial to extend a stereo orthogonal method to use a third photograph in order to improve the reconstruction accuracy. Furthermore, these methods are invariant to the lighting conditions even though inappropriate lighting can spoil the quality of the texture in reconstructed faces. Despite the good features of stereo methods using orthogonal images, these methods create smoothed faces because generic face models used in such cases usually have a relatively small number of vertices that does not allow the geometric description of local facial structures. 3D reconstruction based on orthogonal images can only be used in cases where images are captured in controlled conditions in order to ensure that the input pair of images is indeed orthogonal.

For face reconstruction in CH applications, this method is only applicable for reconstructing faces appearing on statues, since in those cases it is feasible to obtain two orthogonal pictures of the face using ordinary digital cameras. The fact that the texture of the input faces is not utilized for computing the geometry of the face, implies that this method can be applied to statues designed using different materials, colours and tones.

2.2.2 Non-orthogonal Images. Non-orthogonal stereo-based 3D reconstruction is a technique used to 3D reconstruct very diverse scenes. However, as this technique is quite successful, it has also been applied to the 3D face reconstruction domain (Chen, 2001; Leclercq, 2004). Non-orthogonal stereo-based techniques require as input two images of the same object captured from different angles using two calibrated digital cameras. Pixel correspondences are established between the two images to create the disparity map. The disparity map and the knowledge of the relative distance between the two cameras are used to compute the depth map, which shows the geometric structure of the face.

In general, stereo non-orthogonal techniques provide geometrically accurate results with the main drawback being the necessity of structured lighting. In addition, such techniques do not require a face database or a 3D generic face model.

Stereo-based face reconstruction methods can be used for generating 3D face models of statues based on two photographs captured from different angles. This method is also applicable for reconstructing the 3D geometry of faces appearing in CH artefacts in cases where the face of the same subject is shown as seen from different viewpoints. However, reconstructing a 3D model using two unconstrained views of the same subject, possess extra difficulties, as in such cases camera parameters and the distance between the two cameras are not known.

2.3 Video-Based Methods

Video based techniques are used for generating a 3D face model based on multiple face images captured from different view-points. The main steps in video-based 3D face reconstruction are:

1. Capture an image sequence showing the face to be reconstructed as seen from various viewpoints.
2. On the first frame locate a number of key features.
3. Track the location of the features in the remaining frames.
4. Use the location of feature points to reconstruct the geometry of the 3D face.
5. Collect texture information from the video frames and synthesize the texture of the reconstructed 3D face.

Researchers in this area have developed a diverse number of techniques. These techniques share similarities in the initial steps, as they all require appropriate feature point detection and tracking. The 3D face reconstruction step is done using either generic model morphing (Leung, 2000; Liu, 2001), linear combinations of basis functions that do not require a generic model (Bregler, 2000), shape from motion techniques (Chowdhury, 2002) or by using PCA based techniques (Xin, 2005) similar to example-based 3D face reconstruction.

An extension to video-based methods are Silhouette-Based methods. The objective of silhouette based methods is to produce a 3D reconstruction (or optimization) of a face by using multiple face outlines extracted from several face images or a video sequence. A complete set of silhouettes captured from different angles provides details related to the geometrical structure of a face hence it is possible to generate a 3D face from silhouettes. Figure 2 shows typical face silhouettes extracted from an image sequence. Silhouette-based methods have been used either individually or in combination with other 3D reconstruction methods such as stereo (Fua, 1999) and example-based methods (Moghaddam, 2003; Lee, 2003).

Figure 2: Examples of face silhouettes extracted from an image sequence

In applications involving 3D face reconstruction in CH applications, video-based methods are only applicable to the reconstruction of faces appearing in statues, since in such cases it is possible to obtain multiple images of the same face as seen from different angles.

2.4 Image-Based 3D Reconstruction in CH Applications

In the previous section, we have briefly discussed image-based 3D face reconstruction methods belonging to three different categories. In this section we discuss the applicability of those methods to reconstructing images in CH applications.
Among the three categories example-based methods are the most applicable for reconstructing faces from CH artefacts since in such cases only one view of the face is usually available.

However, since typical example-based methods depend on comparing the texture between the model and the face to be reconstructed, these methods are not applicable for reconstructing faces with diverse texture tones and colours such as the ones encountered in CH applications (see figure 1).

Since stereo and video-based methods require two or more views of the face to be reconstructed, such methods are applicable only in the case of statues or in cases that multiple instances of the same face are available on an artefact. Techniques from this category are not so sensitive to the type of texture, hence they can be applied for reconstructing faces with varying textures.

Ideally an image-based 3D reconstruction technique suitable for reconstructing faces appearing in CH artefacts should have the following characteristics:

- Ability to generate a 3D reconstruction from a single image, since certain types of artefacts display faces as seen from a single view.
- Ability to deal with significantly different facial geometries, since artists in different time periods adopt different painting styles, resulting in deformed facial geometries. For example faces showing Saints drawn during the Byzantine period tend to be extremely thin.
- Ability to generate 3D models with textures with different tones, shadings and colours.
- Ability to deal with damaged faces, since in some cases part of the face to be reconstructed is destroyed.

3. 3D FACE RECONSTRUCTION METHOD

In this section we describe a 3D face reconstruction method that can be used for generating 3D models of faces appearing in CH artefacts. In order to deal with most of the issues raised in section 2.4 we developed a dedicated 3D reconstruction algorithm that can be used for generating 3D models of faces shown in CH artefacts. The proposed method is a model-based approach that utilizes a PCA-based 3D shape model (Blanz, 1999) generated during the training process. The main steps involved in the reconstruction process are shown in figure 3. Details about each step of the reconstruction process are provided in the following sections.

3.1 3D Shape Model

The training set consists of 60 3D models of volunteers where each model contains about 75000 vertices. 3D models used in the training set were captured using a laser 3D scanner and pre-processed in order to eliminate noise. Typical samples from the training set are shown in figure 4.

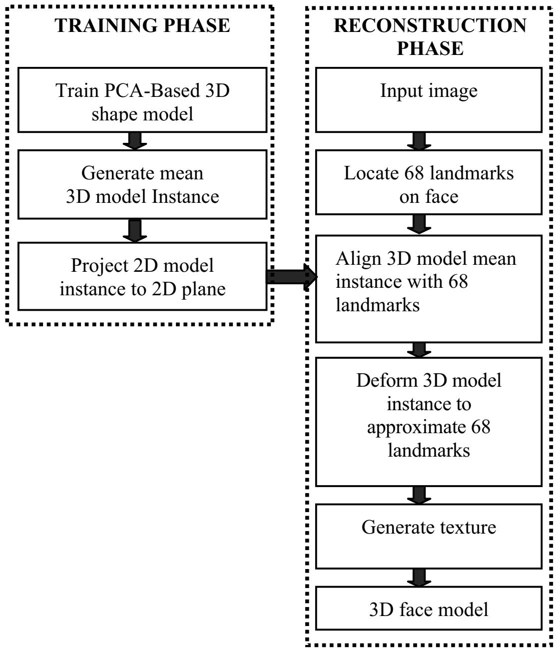


Figure 3: Block Diagram of the 3D reconstruction method

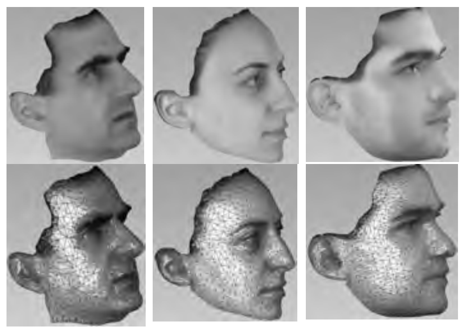


Figure 4: A typical 3D faces from the training set. Row 1: textured models, Row 2: models with triangles overlaid

The model training procedure involves the alignment of all training samples so that there is one-to-one correspondence between vertices of all samples in the training set (Blanz, 1999), and the estimation of the mean 3D face model among the training set. A PCA-based model training procedure similar with the one used by Cootes et al (Cootes, 2001) and Blanz and Vetter (Blanz, 1999) is used for buiding a face shape model. During the process we represent training samples using the coordinates of each vertex and the covariance matrix of the deviations of vertices from the mean shape is estimated. By applying Principal Component Analysis on the covariance matrix the main modes of shape variation within the training set are defined and 3D face shapes from the training set can be represented using a small number of 3D shape model parameters. In this framework model parameters are the weights of the eigenvectors derived during the Principal Component Analysis decomposition process. Based on this framework, it is possible to generate novel 3D face shapes by setting different values to the model parameters. 3D shapes generated by the model display shape instances consistent with the training set. More details related to the process of training such models appear elsewhere (Cootes, 2001; Blanz, 1999).

3.2 3D Face Reconstruction

The 3D reconstruction algorithm involves the following steps.

3.2.1 Location of facial features. Given a face to be reconstructed we locate 68 landmarks on predefined locations as shown in figure 5.

Figure 5: Locations of 68 landmarks on a face to be reconstructed

The location of the landmarks is done based on a supervised semi-automatic method, where the users provide rough estimates of the locations of the face outline and template matching–based and edge-based algorithms are employed in order to locate the correct position for each landmark. Due to the diversity of the texture of faces to be reconstructed, a totally automatic method would lead to inaccurate landmark location. The process of locating the 68 landmarks on a face image, using the semi-automatic tool requires about five minutes to be completed.

3.2.2 3D Model Alignment. During this step we aim to align the 3D model vertices with the 68 landmarks located on the face to be reconstructed. In order to enable the alignment, we first establish the correspondences between the 68 landmarks and the 75000 3D vertices, i.e. among the 75000 vertices we select the 68 vertices that correspond to the locations of the 68 landmarks. The establishment of the correspondences is done once during the training phase using a dedicated visualization tool. Once the 68 vertices from the 3D model are selected, all further processing is carried out using only the selected 68 vertices, minimizing in that way the execution time. Only during the final step of texture mapping all 3D model vertices are used.

During the alignment phase we generate the mean 3D model instance of the 3D shape model (this is done by setting the model parameters to zero) and project the resulting 3D vertices to the 2D plane using an orthographic projection transformation. Among the 75000 vertices projected to 2D we select the 68 that correspond to the landmarks of the face to be reconstructed and calculate the overall distance (dis) between the two sets of points using:

$$dis = \sum_{i=1}^{N} \sqrt{[(\mathbf{X2D}_i - \mathbf{X3D}_i)^2 + (\mathbf{Y2D}_i - \mathbf{Y3D}_i)^2]} \qquad (1)$$

Where N is the number of selected landmarks (i.e 68),

X2D, **Y2D** are the x and y coordinates of the 68 landmarks located on the face to be reconstructed.

X3D, **Y3D** are the x and y coordinates of the selected 68 shape model vertices projected to 2D.

The alignment process involves three steps:

a. *Rigid transformations:* The 2D projected vertices are translated, scaled and rotated around the z-axis until the distance between them and the 68 2D landmarks is minimized. In the actual implementation the Generalized Procroustes Analysis alignment method (Dryden, 1998) is used. This process results in the estimation of the best rigid transformation parameters (i.e. the best x and y translation, scaling factor and z-rotation angle (roll angle)) that ensure the best alignment between the two sets of points.

b. *Definition of Yaw and Pitch angles:* 3D vertices of the shape model are rotated between angles of –60 to 60 degrees around the x and y axis. For each rotation angle the distance dis is calculated (see equation 1) and the combination of the yaw and pitch angles that minimize dis is recorded. In effect during the process the 3D pose of the face to be reconstructed is determined so that projected 3D vertices are better aligned with the 2D landmarks.

c. *Non-rigid transformations:* Up to this point the alignment process is carried out based on transformations of the mean 3D shape. However, for the reconstruction algorithm to function correctly, the mean shape must be deformed in order to approximate the shape of the face in the 2D image. The optimum shape deformation required is estimated by defining the values of the 3D shape model parameters that minimize the distance dis. In this case a minimization algorithm based on a pattern search method (Yin, 2000) is used for defining the optimum values of the model parameters.

The basic steps of the 3D model alignment process are demonstrated in figure 6.

Figure 6: Demonstration of the alignment process. (Blue points are the actual 2D landmarks on the face to be reconstructed and red points are the corresponding 2D-projected vertices of the shape model instance). Left image (Before alignment) (dis=68000). Right image after rigid transformations, adjustment of yaw and pitch angles and adjustment of model parameters (dis=551)

3D Model Generation. The completion of the alignment process results in the definition of transformation and deformation parameters that can be used for aligning the 68 selected 3D vertices on the 68 landmarks located on the face to be reconstructed. The definition of the deformation parameters (i.e the parameters of the 3D shape model) can be used for generating the geometry of the face to be reconstructed. When both the transformation and deformation parameters are applied to all vertices of the 3D model (so far all work was done using the 68 selected vertices) it is trivial to generate the texture of the 3D model by sampling the 2D image at the positions where each vertex is projected. As a result both the geometry and the texture of the face are defined and the 3D reconstruction process is completed.

When the proposed algorithm was implemented in MATLAB (www.mathworks.org) the 3D reconstruction process needs about one minute to be completed. However, this time does not include the time required for locating the 68 landmarks on the face to be reconstructed. About five minutes are also required for locating the 68 landmarks on the face, using the semi-automatic tool.

3.3 Experimental Evaluation

In order to assess quantitatively the accuracy of the proposed method in reconstructing 3D faces, we staged an experimental evaluation. For our experiments we used a dataset containing 100 3D human faces (see figure 4) captured with a 3D laser scanner. We have used 60 3D faces among the dataset for training the system and the remaining 40 3D faces were used for testing the reconstruction accuracy.

During the experiments we train a 3D shape model using the 60 samples from the training set. We then project all 40 faces from the test set to a frontal 2D face image and we define the coordinates of the 68 landmarks located on each test face image. The reconstruction algorithm described in section 3.2 is then used for reconstructing the 3D geometry and texture of the 2D projected faces from the test set. Since for each test face we know the actual geometry and texture of the corresponding 3D face, we obtain quantitative measurements that assess the similarity between reconstructed and actual 3D faces by calculating the Euclidean distance between the actual and reconstructed vertices and color intensities. The results of the experiments are shown in Table 7.

	Mean	Standard Deviation
Shape Error	0.0298 mm	0.0078 mm
Intensity Error	0.15*	0.046*

* When using a 0-255 range of color intensities

Table 7: Quantitative evaluation results

The results of the experiment indicate that the accuracy of reconstructed faces is satisfactory. Bearing in mind that the precision of commercially available 3D scanners is around 0.05mm the performance of the proposed method is deemed adequate for most relevant applications.

3.4 Reconstructing Faces in CH Artefacts

Figure 9 (see page 8 of the paper) shows preliminary reconstruction results when we tested the proposed 3D reconstruction method in reconstructing faces appearing in CH artefacts of different styles and types. The results demonstrate that this method can be used for obtaining plausible reconstructions of the 3D facial appearance of the faces shown in the corresponding artefacts.

4. CONCLUSIONS

In this paper we have provided an overview of image-based 3D face reconstruction methods described in the literature. In particular we described typical example-based, stereo-based and video-based methodologies. In each case we provided a discussion pertaining to the applicability of the methods for reconstructing faces appearing in CH artefacts. According to the relevant discussion a successful method for CH applications should be able to reconstruct a face based on a single view but in order to deal with the diversity of texture variation it should generate the texture of the model through a sampling rather than a generative process.

Based on the conclusions of the review we developed a face reconstruction algorithm, which uses a 3D shape model as the basis of estimating the geometry of a face. During the process, the deformation parameters of a 3D shape model that best approximate the positions of 68 landmarks located on the face to be reconstructed are defined enabling in that way the definition of the 3D geometry of a face. A direct texture mapping approach is then used for generating the texture of the resulting 3D model. The results of a preliminary quantitative performance evaluation prove that the proposed method is able to reconstruct the 3D geometry of frontal faces appearing in images with reasonable accuracy.

Visual results indicate that the proposed method performs well in reconstructing faces appearing in CH artefacts despite the fact that the method was applied to cases involving significant shape and texture variation. In the future we plan to carry out more work in this area in order to deal with the following issues.

- Since the texture is collected directly from the image currently the method is only applicable to reconstructing faces with approximately frontal view in the given image. In the future we plan to upgrade the texture mapping process so that it will be possible to generate the texture of the missing parts based on the visible texture. For this purpose techniques that employ symmetry constraints and related face restoration techniques (Draganova, 2005) will be used. As part of this effort we also plan to augment the basic 3D shape model with 3D models describing the shape of occluding structures that may be encountered in such applications (i.e beards) so that such structures could be added on reconstructed faces. It is anticipated that our overall effort along these lines will lead to efficient and accurate digital restoration techniques that can be applied to damaged faces.
- Currently the reconstruction method relies on the use of a semi-automatic technique for locating 68 landmarks on the face to be reconstructed. In the future we plan to fully automate this process. In order to deal with this issue we plan to use an extended training set that includes faces on different types of artefacts, so that customized templates for specific artefact types will be used during the process of feature location.
- Although the initial quantitative and visual results are promising, we plan to stage a more rigorous quantitative evaluation of the reconstruction accuracy. During the process we plan to reconstruct 3D models of faces of statues and compare the reconstructed models with the actual 3D geometry obtained using 3D laser scanners. We also plan to investigate the sensitivity of the method in reconstructing non-frontal faces.

The ultimate aim of our work is to use the results of our work in different CH related applications. For example we plan to use 3D models obtained using the proposed method for generating complete models that can be used in animations, games and other educational software applications. Also we plan to use the generated 3D models for assessing trends and styles of different artists. For example in figure 8 we show two mosaics of Ktisis (a figure that symbolizes creation) and the corresponding 3D reconstructions. Differences in the face geometry can be used for obtaining conclusions related to the design style adopted in each case.

Figure 8: 3D Reconstructions of Ktisis from two different mosaics. The first column shows the original face and the remaining columns show the reconstructed face as seen from different views. Comparison of the 3D geometry of the two reconstructed models will enable the comparison of different styles and trends

ACKNOWLEDGEMENTS

This work was supported by a Cyprus Research Promotion Foundation research grant (EPYAN/0205/08).

REFERENCES

References from Journals:
Akimoto, T., Suenaga, Y., Wallace, R.S., 1993. Automatic creation of 3D facial models. IEEE Computer Graphics and Applications, 13(5), pp 16-22.

Atick, J.J., Griffin, P.A., Redlich, N.A., 1996. Statistical approach to shape from shading: Reconstruction of 3D face surfaces from single 2d images. Neural Computation. 8, pp. 1321-1340.

Chen, Q., Medioni, G., 2001. Building 3-D human face model from two photographs, Journal of VSLI Signal Processing, 27(1,2), pp. 127-140.

Cootes, T.F., Edwards, G.J., Taylor, C.J., 2001. Active Appearance Models. IEEE Transactions of Pattern Analysis and Machine Intelligence, 23, pp 681-685.

Fua, P., Miccio, P., 1999. Animated heads from ordinary images: A least squares approach. Computer Vision and Image Understanding. 75(3), pp. 247-259.

Liu, Z., Zhang, Z., Jacobs, C., Cohen, M., 2001. Rapid modeling of animated faces from video. The Journal of Visualization and Computer Animation. 12(4), pp. 227-240.

Park, I.K., Zhang, H., Vezhnevets, V., 2005. Image-based 3d face modeling system. EURASIP Journal on Applied Signal Processing. 13, pp. 2072-2090.

Stylianou, G., Lanitis, A., 2008. Image Based 3D Face Reconstruction: A Survey. International Journal of Image and Graphics, (to appear).

Yin, S., Cagan, J. 2000. An Extended Pattern Search Algorithm for Three-dimensional Component Layout. Journal of mechanical design , Vol. 122, No 1, pp. 102-108.

References from Books:
Dryden, L. Mardia, K.V., 1998. Statistical Shape Analysis. John Wiley & Sons.

References from Other Literature:
Blanz, V., Vetter, T., 1999. A morphable model for the synthesis of 3D faces. ACM Siggraph. pp. 187-194.

Bregler, C., Hertzmann, A., Biermann, H., 2000. Recovering non-rigid 3d shape from image streams. Proc. of IEEE Conf. on Computer Vision and Pattern Recognition.

Chowdhury, A.R.K., Chellappa, R., Vo, T., Krishnamurthy, S., 2002. 3D face reconstruction from video using a generic model. Proc of Int. Conf. on Multimedia. pp. 449-452.

Draganova, C., Lanitis, A., Christodoulou, C., 2005. Restoration of Partially Occluded Shapes of Faces Using Neural Networks. Procs. of the IV International Conference on Computer Recognition Systems.

Hu, Y., Jiang, D., Yan, S., Zhang, L., Zhang, H., 2004. Automatic 3D reconstruction for face recognition. IEEE Int. Conf. on Automatic Face and Gesture Recognition.

Leclercq, P., Liu J., Woodward A., Delmas, P., 2004. Which stereo matching algorithm for accurate 3d face creation?, Int. Workshop on Combinatorial Image Analysis, Lecture Notes in Computer Science, 2004.

Lee, W., Kalra, P., Thalmann, N.M., 1997. Model based face reconstruction for animation. Int. Conf. on Multimedia Modeling.

Lee, J., Moghaddam, B., Pfister, H., Machiraju, R., 2003. Silhouette-based 3D face shape recovery. Graphics Interface.

Lee, J., Machiraju, R., Pfister, H., Moghaddam, B., 2005. Estimation of 3D faces and illumination from single photographs using a bilinear illumination model. Proc. of the Eurographics Symposium on Rendering Techniques. pp. 73-82.

Leung, W.H., Tseng, B.L., Shae, Z., Hendriks, F., Chen, T., 2000. Realistic video avatar. IEEE Multimedia and Expo.

Moghaddam, B., Lee, J., Pfister, H., Machiraju, R., 2003. Model-based 3d face capture with shape-from-silhouettes. Analysis and Modeling of Faces and Gestures.

Romdhani, S., 2005. Face Image Analysis using a Multiple Features Fitting Strategy. Ph.D. thesis, University of Basel.

Xin, L., Wang, Q., Tao, J., Tang, X., Tan, T., Shum, H., 2005. Automatic 3D face modeling from video. IEEE Int. Conf. on Computer Vision.

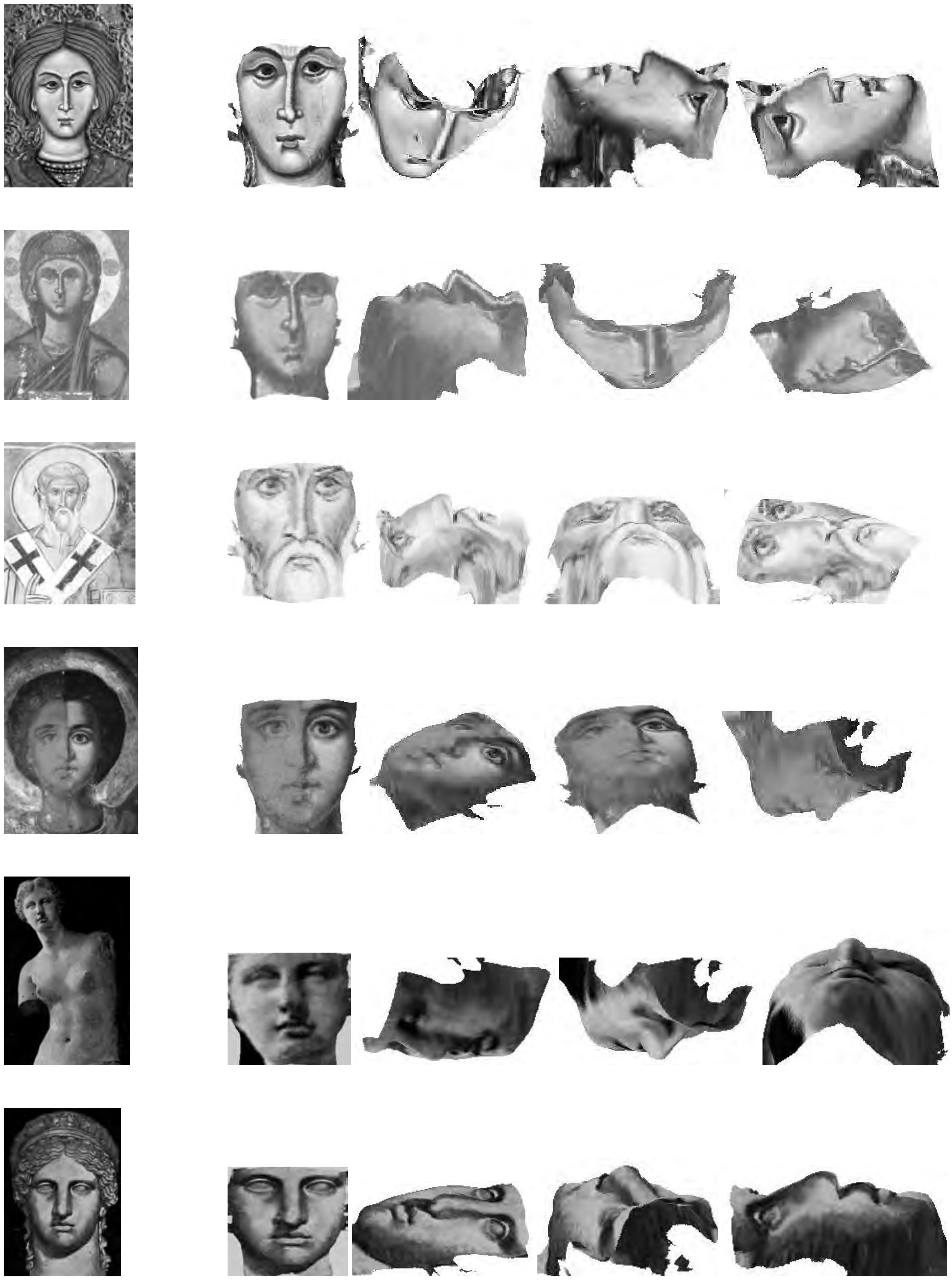


Figure 9: 3D reconstruction of faces appearing in CH artefacts. The first column shows the original face and the remaining columns show the reconstructed face as seen from different views

Spatial Information Management for Cultural Heritage

A GIS IN ANCIENT CARTOGRAPHY: A NEW METHODOLOGY FOR THE ON-LINE ACCESSIBILITY TO THE CARTOGRAPHIC DIGITAL LIBRARIES

P. Chías [a, *], T. Abad [a]

[a] Dept. of Architecture, Technical School of Architecture and Geodesy, University of Alcalá, Spain – pilar.chias@uah.es, tabad@ciccp.es

KEYWORDS: GIS, Cartographic Heritage, Digital Libraries, On-Line Accessibility, Multilingual and Multiformat Databases, Historical Landscape

ABSTRACT:

The frame that set the Council of the European Union about the Digital Libraries has allowed to create important cartographic databases that are accessible on-line and give a response to the real demand among citizens and within the research community. We have developed an open GIS that surpasses the usual operativity of the traditional multiformat databases as it enlarges through the queries the way to get a more personalised information. This new methodology has been created with the aim of being implemented all around the European Union, and will allow the searches and analysis of the historical evolution of the territories and landscapes based upon the study of old cartographic documents.

1. INTRODUCTION

Ancient cartography, as well as old pictures, drawings and photographs, has not been used traditionally as a reliable source of information about the history and the evolution of the land- and the townscape. Those graphic materials have been usually considered as 'second order' documents, mainly because of the difficulties that their interpretation can sometimes involve (Harley, 1968) due to the different conventions that are applied in each case by the cartographer.

But this is not the only reason why cartography is so seldom used in the historical searches, because there are other problems related to the difficulties of their localisation and visualisation that have to be considered.

Obviously, it is not easy to access to an original big size and small-scale map that is sometimes composed by several printed sheets; and it is also complicate to see properly the symbols employed in the map and read its texts when it is imposed to handle a reduced hardcopy or a low resolution digital image.

Although we find it is not essential to have an exhaustive knowledge of the context of each map to get a meaningful interpretation of it (Skelton, 1965: 28; Andrews, 2005), it is necessary to achieve some basic specific concepts on the theory of the cartographic expression and design (about map projections, symbols or representation of relief, for instance), because the lack of them can difficult the right interpretation of the document and twist the results of the investigations (Vázquez Maure and Martín López, 1989: 1-10).

According to the strategies of the Council of the European Union about the European Digital Libraries considered as a common multilingual access point to Europe's digital cultural heritage, and assuming that the ancient maps and plans are important cultural materials, in the last decade there have been created several cartographic databases that allow an efficient on-line accessibility.

Although the main Spanish cultural offices are making a strong effort to digitise the public collections of historical documents, the problems posed by the different locations, techniques, sizes and preservation conditions, are delaying the prompt achievement of their diffusion. And we have to mention another problem that is associated to the difficulties of finding those maps, because they are frequently included into other documents or inside bundles of old papers, and remain yet undiscovered.

Among those important initiatives we will emphasize the one created by the Institut Cartogràfic de Catalunya and those of the Portal de Archivos Españoles (PARES), that not only show low-resolution images of each map, but provide an accurate description of the document and the conditions of use.

There are also other experiences that integrate old maps into GIS as the *Gregoriano Cadastre* (Orciani et al., 2007) and the old cadastral maps of Utrecht (Heere, 2006), but they are focused on a deep knowledge of the reconstruction of the old properties.

With the aim of widening the above mentioned strategies of the Council of the European Union, we have developed an innovative GIS based methodology in ancient cartographic documents, that allows to get the information from the relational cartographic databases that we have elaborated not only through the traditional queries, but applying the hypermedial concept.

But the final target of this project is to diffuse the old cartographic treasures that compose a relevant part of the Spanish cultural heritage, that actually remains unknown to the public and even to a great number of specialists (Chías and Abad, 2006; 2008).

2. THE CARTOGRAPHIC DATABASES

2.1 Contents and Structure

To define the contents of our cartographic database we have decided to apply the concept of cartographic document in the wide sense of Harvey (1980: 7) and Harley and Woodward (1987: 1, xvi), that includes all kinds of maps, plans and charts at different scales (architectural, urban and territorial scales), as well as pictures and bird's-eye views (Kagan, 1986: 18-26; De Seta, 1996), with no restrictions due to techniques, functions or origins (Fig. 1).

As we must also restrict the temporal and the geographical subject of the contents of the cartographic databases, we firstly decided to include all the historic documents that have been drawn before 1900, mainly because along the 20th century the cartographic production and techniques have very much increased in many senses and its study should be carried separately. Secondly, the spatial restriction has been imposed to the search and we decided that the cartographic database should concern the actual Spanish territories (Chías and Abad, 2008 a).

Figure 1: José Pastrana, Escribanía de Zarandona y Balboa, cadastral map [*Carta topográfica de los términos de Villagonzalo-Pedernales y Renuncio, separados por dos hileras de mojones*], 1758 (Valladolid, Archivo de la Real Chancillería, Planos y Dibujos, Óleos 0014).

Assuming all those circumstances, the former stages of our search have focused on finding, studying and cataloguing all kind of cartographic documents that are preserved in the main Spanish collections, archives and libraries as the Biblioteca Nacional de España (Madrid), Biblioteca del Palacio Real (Madrid), Biblioteca del Monasterio de El Escorial (Madrid), Biblioteca de la Universidad de Barcelona, Biblioteca de la Universidad de Salamanca, Real Academia de la Historia (Madrid), Real Academia de Bellas Artes de San Fernando (Madrid), Archivo Histórico Nacional (Madrid), Archivo General de Simancas (Valladolid), Archivo de la Real Chancillería (Valladolid), Archivo del Centro Geográfico del Ejército (Madrid), Instituto de Historia y Cultura Militar (Madrid), Instituto Geográfico Nacional (Madrid), Museo Naval (Madrid) and Archivo de Viso del Marqués (Ciudad Real) among many others, as well as local and ecclesiastical archives.

As digital preservation implies copying and migration, it has always been considered in the light of IPR legislation (Commission of the European Communities, 2005): the digitised funds of other libraries are precisely quoted and respect the conditions that have been established for consulting the documents by the rightholders. We also provide the links in case that the documents are included in other cartographic databases that can be accessed through Internet.

To digitise the non-digitised funds we apply both the contact scanning and the non-contact photographic methods (Tsioukas, Daniil and Livieratos, 2006), trying to minimize the distortion problems by digitising each sheet separately, although this process can not eliminate other problems in the final assemblage of the mosaic image.

The maps that are yet unpublished or remain unknown (for instance, those of the private collectors) if there is any legal obstacle that allows them to be shown, we offer a link to a high resolution image 1:1 that makes possible to see every detail and to read every name to analyse it properly.

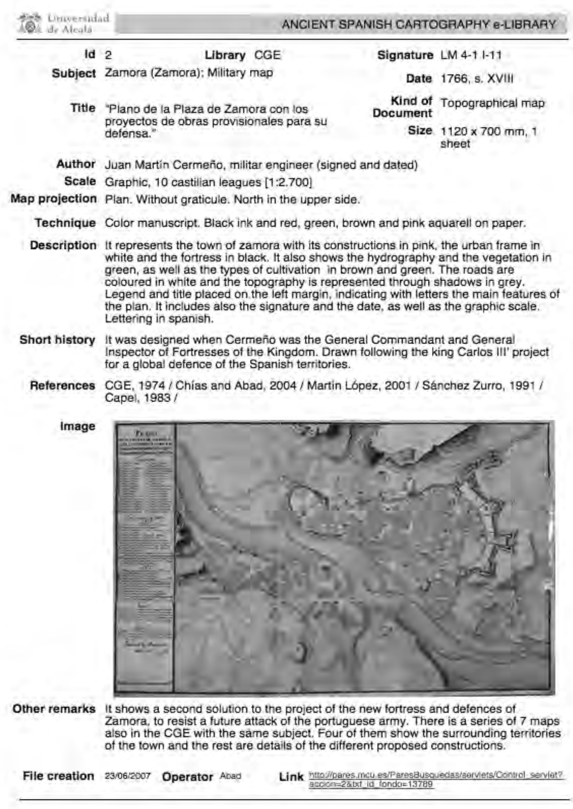
The next step was to construct the relational databases over a commercial compatible platform. They have been designed as multilingual (there is an English version) and open ones to allow including new registers in the future and even adding new fields or tables, to update the contents to the new needs without damaging the existing ones. Moreover the concept 'relational' implies the possibility of crossing the data of the different tables and reducing their weights, making easier the data management and the queries.

According to this, our methodology includes three main tables, that are the following:

- 'Cartography', that contents all the registers concerning the cartographic documents and follows the ISBD Norms of cataloguing.
- 'Bibliography', that includes the complete bibliographical references that appear in the field *Bibliography* of the table 'Cartography'.
- 'Libraries, Archives and Map Collections', is the table that includes the complete references of the collections that have been visited, and that appear just as an acronym in both *Collection* and *Signature* fields of the table 'Cartography'.

The three tables have been designed sharing at least one field that allows crossing the datafiles and economizes data length in the databases.

The design of the table 'Cartography' joins both the descriptive and the technical data about each document, joining the perspectives of the historian and the cartographer. The items that have been included are the following (Fig. 2):



Universidad de Alcalá — ANCIENT SPANISH CARTOGRAPHY e-LIBRARY

Id	2 **Library** CGE	**Signature**	LM 4-1 I-11
Subject	Zamora (Zamora): Military map	**Date**	1766, s. XVIII
Title	"Plano de la Plaza de Zamora con los proyectos de obras provisionales para su defensa."	**Kind of Document**	Topographical map
		Size	1120 x 700 mm, 1 sheet
Author	Juan Martín Cermeño, militar engineer (signed and dated)		
Scale	Graphic, 10 castilian leagues [1:2.700]		
Map projection	Plan. Without graticule. North in the upper side.		
Technique	Color manuscript. Black ink and red, green, brown and pink aquarell on paper.		
Description	It represents the town of zamora with its constructions in pink, the urban frame in white and the fortress in black. It also shows the hydrography and the vegetation in green, as well as the types of cultivation in brown and green. The roads are coloured in white and the topography is represented through shadows in grey. Legend and title placed on the left margin, indicating with letters the main features of the plan. It includes also the signature and the date, as well as the graphic scale. Lettering in spanish.		
Short history	It was designed when Cermeño was the General Commandant and General Inspector of Fortresses of the Kingdom. Drawn following the king Carlos III' project for a global defence of the Spanish territories.		
References	CGE, 1974 / Chías and Abad, 2004 / Martín López, 2001 / Sánchez Zurro, 1991 / Capel, 1983 /		
Image			
Other remarks	It shows a second solution to the project of the new fortress and defences of Zamora, to resist a future attack of the portuguese army. There is a series of 7 maps also in the CGE with the same subject. Four of them show the surrounding territories of the town and the rest are details of the different proposed constructions.		
File creation	23/06/2007 **Operator** Abad	**Link**	http://pares.mcu.es/ParesBusquedas/servlets/Control_servlet?accion=2&txt_id_fondo=13789

Figure 2. One of the filling cards.

- *Place or Subject* (text field): Refers to the geographical place that is represented in the document and the province it belonged

to, since the reform of 1831. To define clearly the territorial limits, the old councils or boundaries are also included. And to determine the original uses of the map, it is also specified if it is a general, thematic (geological, military, statistic, cadastral, etc.) or a topographical map or plan.
- *Date* (numerical field): as precise as the document can be dated. If it is only an approximated date, it is quoted among square brackets.
- *Kind of Cartographic Document* (text field): it defines if it is a map or plan, a chart, a portolan or a view, or even a terrestrial globe.
- *Size* (text): width and high of the image in mm; it is also included the total size of the sheet(s) (or other supporting materials) and the number of pieces or sheets that compose the ensemble of the document.
- *Collection* and *Signature* (text): the collection that preserves the document and the signature; the first one is quoted through an acronym and the second one is abbreviated according to the norms (its total extension can be consulted in the table 'Libraries, Archives and Museums'); if possible, it includes the link to other e-libraries or references.
- *Original Title* (text): quoted among inverted commas if it is literal, as it is written in the document; otherwise it is defined among square brackets through its main features.
- *Author(s)* (text): names of the author(s) if the map is signed; in case of ascription, the name(s) appear among square brackets; they can be also 'unknown'.
- *Scale* (text): it is defined graphically or as a fraction, detailing the different units employed; when there is no definition scale, appears 'without scale'.
- *Projection* (text): details the projection employed with its different elements: grid, references, orientation; it is also referenced the use of different projections, for instance profile or section added, axonometrics or perspective views, with their own distinctive elements, and even the case of large scale plans.
- *Technique* (text): makes a distinction between manuscripts and printed maps, as well as the drawing surfaces and techniques, specifying the uses of colour.
- *Short History* (of the map, text): place of edition, editor, or if it is a part of a big compilation or atlas. It is also mentioned where it comes from or the precedent owners, and the date of purchasing.
- *References* (of the map, text): abbreviated and following the international system for scientific quotations ISO 690-1987.
- *Image* (object/container field): it is included a low resolution raster image in highly compressed jpg format of the cartographic document. By clicking on the image, it is possible to display a high resolution one in a *tiff* format that allows to see the details and to read the texts. If the map is composed by several sheets, it is possible to see each one separately (and to compose it apart).
- *Other Remarks* (text): in the case of a printed map, includes other collections that have a copy, or variations of the plate, as well as manuscript notes, etc.
- *Date* (of the catalogue, autom.).
- *Operator* (for future updates).
The table 'References' defines completely the abbreviations and acronyms used in the other tables. The quotations follow the ISO 690-1987 Norm. The fields are in this case (Fig. 3):
- *Author(s)* (text).
- *Date* (of edition, numerical).
- *Title* (of the book, text).
- *Article's Title* (text).
- *Periodical or Book* (in case of articles or book chapters, text or link).
- *Publisher* (text).
- *Place* (text).
- *Volume* (numerical).
- *Pages* (text).
- *Quotation* (text): as it appears in the other tables.

REFERENCES

Author(s)	Vázquez Maure, F.	**Quotation**	Vázquez Maure, 1982
		Date	1982
Title			
Article's Title	"Cartografía de la Península: siglos XVI a XVIII"		
Periodical or Book	*Historia de la cartografía española*		
Publisher	Real Academia de Ciencias Exactas, Físicas y Naturales		
Place of Edition	Madrid	**Volume**	**Pages** 59-74

Figure 3: The contents of the Table 'References'

Finally the table 'Libraries, Archives and Map Collections' makes possible to identify the acronyms used in the *Collection and Signature* field. It contains the following fields that complete the location of the documents (Fig. 4):
- *Library (Abbreviation)* (text).
- *Collection (Extended Name)* (text).

LIBRARIES, ARCHIVES AND MAP COLLECTIONS

Library (Abbreviation)	IHCM
Extended Name	Instituto de Historia y Cultura Militar, Archivo General Militar, Madrid (former Servicio Histórico Militar)

Figure 4: The contents of the table 'Libraries, Archives and Map Collections'

3. THE GIS IN ANCIENT CARTOGRAPHY

3.1 The cartographic base

The GIS is supported by a commercial platform (GeoGraphics©) that includes a complete and easy to use computer-aided mapping module in a vector format (MicroStation©) and that is a standard with the maximal compatibility, although the connection with the database (ACCESS©, FileMaker©) must be established through an ODBC protocol.
We have tried some other possible GIS platforms that:
- Integrated both the databases and the computer-aided mapping (ArcInfo©).
- Or imported the cartography in and exchange and export format such as the *dxf* (MapInfo©).

The fist one were the most complete of all, but the databases were not so easy to manage and it was sometimes difficult to make changes in their structures.
The second group of GIS platforms allows an easy and flexible management of the databases, but the problems appear as the cartographic base must be imported, because the exchange formats always suppose a loose of the information and the integrity of the graphic elements that leads to a further hard process of validation of the digital vectorial cartography (Chías, 2004; 2004 a).

Figure 5: The cartographic base in a vector format

The possibility of drawing our own vector cartography inside the GIS not only avoids the problems derived from the import of graphic files through the *dxf* format, but allows an easy definition of the base maps, that only need to make a clear definition of the graphic features that must be digitised and of the strategies that must be followed to compose the base maps (Fig. 5).
We have chosen not to use the existing digital cartography that is available because it would suppose a hard process to clean, verify and structure it topologically. We found it would be harder to extract and add the features that were interesting to our GIS in the existing cartography, that to draw a new one specifically designed for it.
The new cartographic base has been structured in data layers and sheets at a 1:200.000 scale, and we have selected the graphic formats, the georeference and the symbolism according to the guidelines of the Instituto Geográfico Nacional of Spain, that ensures a proper understanding of both topographic and planimetric data as well as an easy connection to other existing or future GIS.

3.2 The connection of the digital vector cartography and the databases

Due to the election we made of the GIS platform, the connection of both cartographic, graphic and attribute data storages is made through an ODBC protocol (Fig. 6).
Both ACCESS© and FileMaker© are easy to handle and don't need to give a special training to the operators, but the first one makes 'heavier' databases than the second one, and this is a disadvantage when managing big amounts of data, as we do. FileMaker brings also another important advantage, because the fields of the tables are not limited to 256 characters, and this allows to introduce more information if required.
Each ancient map is referred to a point, a line or a centroid, depending on the territory they represent. Points are used generally in town plans or views; lines are used fundamentally to identify roads, railways and other lineal human-made structures. Centroids are mainly used to identify the maps that represent the different territories.
Obviously, a single centroid can be related to more than a map, and the different possibilities are simultaneously shown in a first screen; afterwards the different documents can be chosen separately.
On the other hand, the contours of the different documents are drawn and referred to their own centroids: then it is quite easy to see the way the maps and plan overlap each other, what allows to perceive rapidly the portions of land that are covered by different maps and which are the historical periods when they have focused the attention due to different targets.

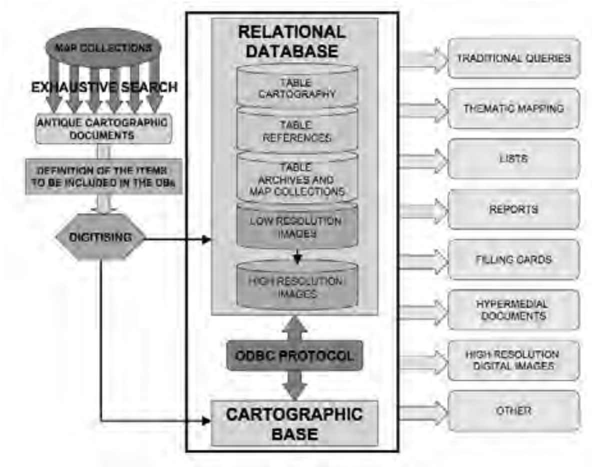


Figure 6: Schema of the GIS structure

As it is shown, we have implemented three main datasets: the cartographic base, the tables that contain the features that we have considered as relevant to define each ancient map, and the sets of images, that are divided in high and low-resolution images of the map.
The datasets of the ancient maps include already more than 3.000 files, that are continuously being updated and distributed in the servers by geographical units.
Obviously, the filling cards access directly to the low resolution image sets, to prevent problems on handling the different tables if they become too 'heavy'.
Only in exceptional cases, and only when the copyright owner allows it, the high resolution images are available.
This system and structure has been the most useful and fits with the strategies that we have previously planned to get the most varied and personalized information of the GIS.

4. THE ON-LINE DIFFUSION THROUGH INTERNET

4.1 Much more than queries

Nowadays the web-based digital resources are quite frequent as a way to preserve and diffuse the cartographic heritage as well as to access to the modern cartography (Zentai 2006; Livieratos 2008).
Previous experiences as the one implemented on the Greek region of Macedonia (Jessop, 2006) or the GIS-Dufour (Egil and Flury, 2007) have shown the potential of GIS and its accessibility through the web.
Our methodology has increased the possibilities of the usual queries that a GIS brings, just formulating them to the different databases separately or even crossing them; but to ensure the proper display of the cartographical information we have designed a filling card as one of the main printable output ways.
Obviously, traditional outputs as thematic mapping, statistics, lists or reports are also available (Spence, 2007; Tufte, 2005) (Fig. 7).

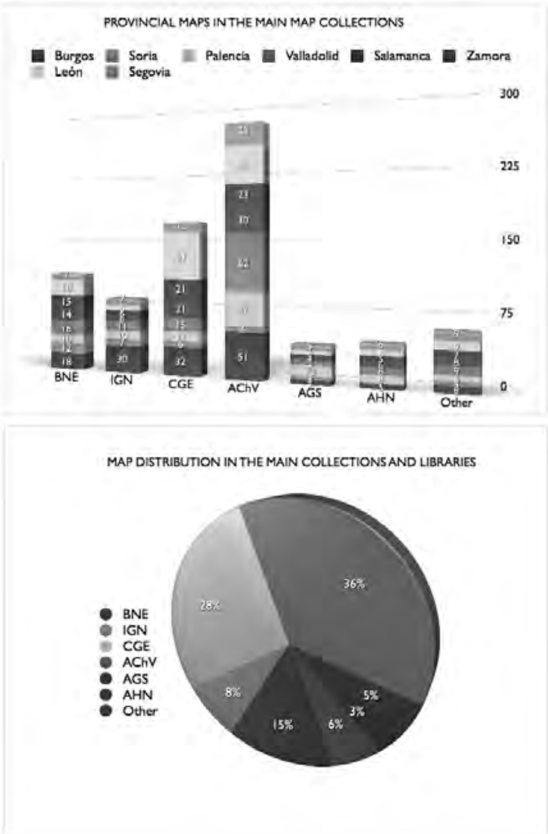


Figure 7: Some statistics obtained through GIS (searches limited to the Castilla-León zone)

This filling card includes also the adequate links to access to other e-libraries or references as it has been above mentioned.
The possibilities that the hypermedial concept brings on getting personalised information of the different data sets are an added value to the traditional queries system.
And as the IPR legislation has been considered from the starting point, in this case we must neither set other supplementary caution that restrict the access to the different data sets, nor establish different access levels.

5. CONCLUSIONS

According to the initiative of the Council of the European Union about the European Digital Libraries as a common multilingual access point to Europe's digital cultural heritage, and considering the ancient maps and plans as important cultural materials, we have developed an innovative GIS based methodology in ancient cartographic documents, whose essential values are:

- To create new cartographic relational, multiformat and multilingual databases that organise and unify the information that different archives and libraries have elaborated about their different funds, as well as to incorporate the dispersed and unknown documents that belong to non-digitised collections. This new information follows the ISBD Norm, and joins and completes the different approaches of the librarian, the historian (Edney, 2007) and the more technical of the cartographer.
    - The new databases join both already digitised materials as well as new information that we have directly produced in a digital format. These circumstances allowed us to get some mechanisms that facilitate the digitalisation of maps, to identify problems and to monitor bottle-necks (as those that appear handling big size maps).
    - They allow also to preserve the original materials, that are usually fragile.
- The open GIS surpasses the usual operativity of the traditional multiformat databases as it enlarges through the queries the way to access to the different kind of data. But we have also disposed a new and personalised way to access to high resolution digital images of the documents by applying the hypermedial concept.
- On-line accessibility and diffusion through the Internet, as a response to a real demand among citizens and within the research community, always paying attention to the full respect to the international legislation in the field of intellectual property.
- This new methodology has been created with the aim of being an open one that allows being implemented in all the countries of the European Union.

The pilot experience about the Spanish ancient cartography that we present is also an example of the full application and the success of the methodology.

6. REFERENCES

Andrews, J.H., 2005. Meaning, knowledge and power in the philosophy of maps. In J.B. Harley, *The new nature of maps. Essays in the history of cartography.* Baltimore, Maryland: The Johns Hopkins University Press, pp. 21-58.

Chías. P., 2004. *Bases de datos y gestores de bases de datos para los sistemas de información geográfica.* Publicaciones de la Escuela Técnica Superior de Arquitectura, Universidad Politécnica de Madrid.

Chías, P. 2004 a. *La imagen de los fenómenos geográficos en un sistema de información geográfica.* Publicaciones de la Escuela Técnica Superior de Arquitectura, Universidad Politécnica de Madrid.

Chías, P. and T. Abad, 2006. A GIS in Cultural Heritage based upon multiformat databases and hypermedial personalized queries. *ISPRS Archives* XXXVI-5, pp. 222-226.

Chías, P. and T. Abad, 2008. Las vías de comunicación en la cartografía histórica de la cuenca del Duero: construcción del territorio y paisaje. *Ingeniería Civil* 149, pp. 79-91.

Chías, P. and T. Abad, 2008 a. Visualising Ancient Maps as Cultural Heritage: A Relational Database of the Spanish Ancient Cartography. In *Information Visualisation*. London, IEE, pp. 453-457.

Commission of the European Communities, 2005. *i2010: Digital Libraries.* Brussels.

De Seta, C. (ed.), 1996. *Cittá d'Europa. Iconografia e vedutismo dal XV al XIX secolo.* Napoli.

Edney, M.H., 2007. Maps and "other akward materials": critical reflections on nature and purpose of cartobibliography. In: *Paper and poster abstracts of the 22nd International Conference on the History of Cartography,* Berne, p. 119.

Egil, H.-R. and Ph. Flury, 2007. GIS-Dufour: historical maps as base in a geographical information system. In: *Paper and poster* *abstracts of the 22nd International Conference on the History of Cartography,* Berne, p. 196.

Harley, J.B., 1968. The evaluation of early maps: Towards a methodology. *Imago Mundi* 22, pp. 68-70.

Harley, J.B. and D. Woodward, 1987. Preface. In J.B. Harley and D. Woodward, *The history of cartography: Cartography in prehistoric, ancient, and medieval Europe and the Mediterranean.* Chicago, Ill.: The University of Chicago Press, Vol. I, pp. xv-xxi.

Harvey, P.D.A., 1980. *Topographical maps. Symbols, pictures and surveys.* London: Thames and Hudson.

Heere, E., 2006. The use of GIS with property maps. *e-Perimetron* 4 (1), pp. 297-307.

Jessop, M., 2006. Promoting cartographic heritage via digital resources on the Web. *e-Perimetron* 3 (1), pp. 246-252.

Kagan, R.L., 1998. *Imágenes urbanas del mundo hispánico, 1493-1780.* Madrid: Eds. El Viso.

Livieratos, E. 2008. The challenges of Cartographic Heritage in the Digital World. In: *Third International Workshop 'Digital Approaches to Cartographic Heritage.* Barcelona, Introductory Speech. http://web.auth.gr/xeee/ICA-Heritage/Commission/3rd_Workshop/Livieratos_Heritage_Barcelona.pdf (accessed 20 july 2008)

Orciani, M., V. Frazzica, L. Colosi and F. Galletti, 2007. Gregoriano cadastre: transformation of old maps into Geographical Information System and their contribution in terms of acquisition, processing and communication of historical data. *e-Perimetron* 2 (2), pp. 92-104.

Skelton, R.A., 1965. *Looking at an early map.* Lawrence, Kansas: University of Kansas Library.

Spence, R. 2007. *Information Visualization. Design for Interaction.* Pearson, London.

Tsioukas, V., M. Daniil and E. Livieratos, 2006. Possibilities and problemas in close range non-contact 1:1 digitization of antique maps. *e-Perimetron* 3 (1), pp. 230-238.

Tufte, E.R., 2005. *Visual Explanations. Images and Quantities, Evidence and Narrative.* Graphics Press, Cheshire, Connecticut.

Vázquez Maure, F. and J. Martín López, 1989. *Lectura de mapas.* Madrid: Instituto Geográfico Nacional.

Zentai, L., 2006. Preservation of modern cartographic products. *e-Perimetron* 4 (1), pp. 308-313.

7. ACKNOWLEDGEMENTS

This paper is a result of two main searches:
- The Project EH-2007-001-00 "Las vías de comunicación en la cartografía histórica de la Cuenca del Duero: Construcción del territorio y paisaje" that has been financed by the Centro de Estudios Históricos de Obras Públicas y Urbanismo (CEHOPU-CEDEX) of the Ministerio de Fomento (Spain).
- The Project PAI08-0216-9574 "La cartografía histórica de la Comunidad de Castilla-La Mancha en los principales archivos españoles", financed by the Consejería de Educación y Ciencia de la Junta de Comunidades de Castilla-La Mancha.
Both are inscribed into our searchers' guidelines on the investigation of the cultural heritage through the application of the most innovative technologies, as GIS and multiformat databases, that set up an essential basis for the knowledge of the history of the territory, the landscape and the town. Since a decade our team is engaged on setting up different useful methodologies that are being implemented in the Technical School of Architecture and Geodesy of the University of Alcalá.

REVEALING THE HERITAGE: CONSTRUCTING A GEOPROCESSED DATABASE FOR MONITORING THE ARCHITECTURAL PATRIMONY IN RIO GRANDE DO NORTE/BRAZIL

E. B. F. Trigueiro [a], V. A. S. de Medeiros[b], A. C. S. Oliveira[c]

[a] Associate Professor, UFRN/CT – Departamento de Arquitetura, Campus Universitário Lagoa Nova, Natal – RN, Brazil, CEP 59072-970, Phone: +55 (0) 84 3222-6577, e-mail : edja.trigueiro@oi.com.br
[b] Associate Researcher/Universidade de Brasília, Architect/Chamber of Deputies – SQN 406, Bloco I, Apto. 202, Asa Norte, Brasília – DF, Brazil, CEP 70847-090, Phone: +55 (0) 61 3349-6798, e-mail: vaugusto@digi.com.br
[c] Architectural Student, (CNPq Grant for Scientific Initiation), UFRN/CT – Departamento de Arquitetura, Campus Universitário Lagoa Nova, Natal – RN, Brazil, CEP 59072-970, Phone: +55 (0) 84 3217-9677, e-mail : alanacso@gmail.com

KEY WORDS: Geoprocessed Database, Built Patrimony, Architectural Heritage, Older Towns, Rio Grande do Norte/Brazil

ABSTRACT:

The article presents procedures and criteria for the construction of a geoprocessed database, an internet website and an Atlas of Architectural Heritage about the built patrimony in the state of Rio Grande do Norte (RN), Brazil. The database articulates information resulting from studies developed by MUsA (Morfologia e Usos da Arquitetura) researchers who study morphology and uses of architecture at UFRN (Universidade Federal do Rio Grande do Norte). It explores possibilities offered by GIS and ICT tools for building analytical resources and tools for the rehabilitation of old town centres, the preservation of architectural heritage, and the disseminating of knowledge about our built patrimony. Studies are based on the representation, quantification, articulation and correlation of variables that must be considered for a broader understanding of the formation and transformation processes of historically important buildings and urban sites. The assumption underlying these procedures is that information, which may be spatialised, become much more intelligible when anchored to a system with a high potential for easy and quick data articulation and visualisation (GIS/ internet), offering resources for fostering comprehension (and perhaps sympathy) concerning the remains of a legacy from past generations, which may, or may not, reach the future ones.

1. RIO GRANDE DO NORTE'S BUILT HERITAGE

The state of Rio Grande do Norte (RN), in the Northeast of Brazil, has, for various reasons beyond the scope of the present paper, lagged behind other Brazilian states as the object of heritage conservation actions. Only 14 buildings in the state are listed by the national heritage agency – the *Instituto Histórico e Artístico Nacional* - IPHAN (National Institute for Historic and Cultural Preservation) – against over 200 in the state of Minas Gerais, where most well preserved 18th century ensembles in the country are concentrated, and over 20 in the adjacent state of Paraíba, a state of comparable size, time of occupation, historical and economical relevance.

This circumstance is aggravated by the fact that knowledge concerning the existing built patrimony is still far from satisfactory both in terms of coverage and systematization. A steady increase in architectural heritage studies (mostly developed at the Universidade Federal do Rio Grande do Norte - UFRN, principal institution of research and higher education in the state) is helping to reduce this gap. In spite of that, most people outside the academic community remain largely ignorant of and indifferent to the destinies of older buildings.

Within this framework the UFRN research group in morphology and the use of architecture (MUsA - Morfologia e Usos da Arquitetura) have been conducting building inventories in older districts of Natal (the state capital city and its first urban settlement) and in some of the oldest hinterland town centres in RN. These inventories are then structured into databases in which digital media is utilised as resources for: (1) expanding the knowledge about our local heritage, especially that which is rapidly fading away in the absence of official conservation actions; (2) supporting decisions in planning interventions that include areas where older buildings remain; (3) improving comprehension about the dynamics underlying the transformation of the built environment, which exerts strong effects over patterns of land use and may determine the substitution of the existing building ensemble; (4) establishing methodological routines and providing tools and data for a range of studies that focus on the formation and transformation of the built environment in RN; and (5) raising awareness about the built heritage beyond the boundaries of academia, by divulging results through freely accessible media.

This paper presents two examples of databases developed at MUsA, in order to contribute towards some of the aims cited above. The first focus on Ribeira, one of the oldest urban settlements of Natal, which is being discussed as a target for redevelopment since 1994; the other records the architectural legacy of centuries of occupation, which survived until the turn of the millennium in Seridó, a region that concentrates some of the oldest towns in RN (figure 1). Both databases address items (1), (4) and (5) as they serve to expand knowledge and awareness about our local heritage and its vulnerability, and offer analytical tools and information for studies concerning the built environment of RN in a broader perspective. The Ribeira database is also an instance of application of digital technology to support planning actions. It was structured to support decisions for the urban rehabilitation plan currently being conceived as it specifically includes variables thought to enhance understanding about the transformation process that led to the area's present state and to provide alternatives for its rehabilitation.

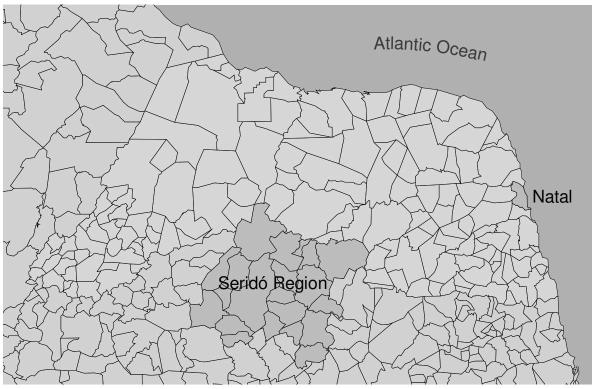


Figure 1: Natal and Seridó in RN, Brasil

2. RIBEIRA

2.1 Towards a Rehabilitation Plan: about what is there and how it came to that

The district of Ribeira, located in Natal, (figure 2) capital city of the state of Rio Grande do Norte, Brazil, sites a fairly robust stock of old buildings.The area expanded from the town's foundation site towards the Potengi river banks, in the beginning of the 19th century, due to requirements of a harbour built to respond to commercial needs.

Ribeira was the busy town core (meaning roughly what would be referred today as a Centre Busyness District) during the early years of civil aviation – when Natal was a mandatory stop for transatlantic flights between Europe and Brazil – and in World War II - when Natal sited an important base for the North American military forces. From the 1970s on, Ribeira suffered a continuing process of decay, which although reducing its vitality and resident population was also partly responsible for the permanence of the architectural stock still existing today, albeit in a degraded state of preservation and conservation.

Among the contemporary actions for the redevelopment of that district, the PRAC-Ribeira – a federal and local government jointly-funded plan for the rehabilitation of urban central areas (*Plano de Reabilitação de Áreas Urbanas Centrais*) – has been the most comprehensive effort to date. In this Plan, inclusive housing, social and functional diversity, and the enhancement of a cultural identity are considered key factors for moulding the area into a sustainable historic centre. Guidelines for actions concerning some physical attributes of buildings and open spaces which may contribute to promote those factors were outlined as part of an urban design proposal, one of the subprojects that is expected to be articulated with others, such as the transportation re-engineering. This proposal, largely underpinned by information that constitute the Ribeira database – entitled Real Estate Database – includes the redefinition of the Preservation Area perimeter, a set of directions for intensifying specific uses in some sectors, and the rehabilitation, re-use or substitution of public and private buildings, particularly empty and partially empty ones, with a special focus on housing.

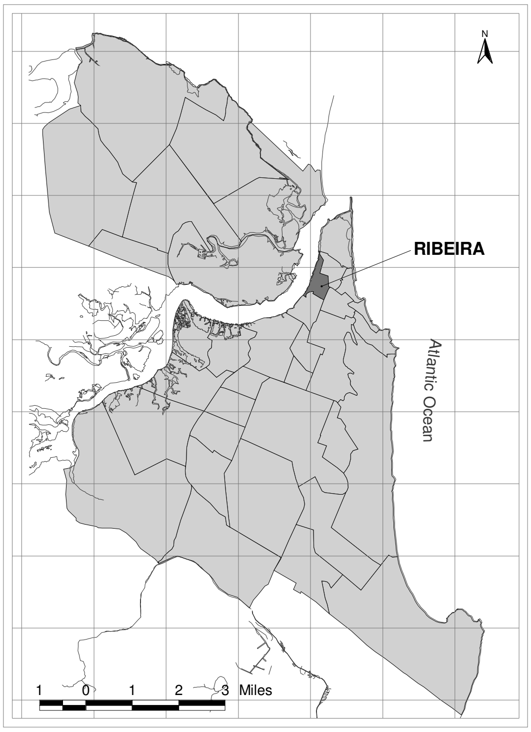


Figure 2: The district of Ribeira in Natal

2.2 Ribeira's Real Estate Database: building location and attributes

The database focus, therefore, on the physical features and current uses of all edifices in the district and defines a set of buildings and groups of buildings whose inherent qualities grant them grounds for distinct categories of recommendation. The articulation of some surveyed variables considered important for the understanding of the potentials of the neighbourhood to meet the PRAC/Ribeira's proposed aims resulted in a set of thematic maps or information plans which illustrate this text.

Photographic images and information were obtained on location by means of field survey, and subsequently structured into a geoprocessed database supported by a geographic information system (GIS). The utilised computer application (SPRING) is a freeware created and made available by the national institute of space research – INPE (*Instituto Nacional de Pesquisas Espaciais*). The database synthesizes information recorded in the field survey regarding location – address and number of the building – plus sketches representing: the position of the building in the plot, its front and side limits visible from the street; and the position of the plot within the block.

For defining the architectural character of each building, the recorded variables were: stylistic affiliation, constructive nature, number of storeys, state of conservation and preservation

Here the term "conservation" refers to the integrity of the building physical structure, which can be: (a) Precarious, when damages risking the stability of the building can be observed.(i.e. absence of roofing elements or doors and windows, resulting in internal exposure to the weather); (b) Regular, when the damages do not (yet) risk the integrity of the building but may lead to this, if not repaired; and (c) Good, when there are no external evidences of damages that may cause risk to the physical integrity of the building, at short or medium term. In addition, listed buildings were recorded, as well as those included in planned or ongoing refurbishment or conservation actions.

The term "preservation" refers to the maintenance of the formal features of the building from the time when it was erected or suffered some thorough stylistic updating intervention. A building is considered: (a) Preserved, when there is stylistic coherence between the elements composing the façade or façades that are visible from the public space; (b) Modified, when the altered elements can be identified (i.e. a reshaped doorway) so that the façade may be restored to its original composition without requiring architectural prospection; and (c) Disfigured, when modifications were so severe that its original features can no longer be identified, even though some of their vestiges are still visible (i.e. a remaining parapet above the façade replaced by large glass/metal doors or shop windows).

2.3 Ribeira's Real Estate Database: on the use of the buildings

Detailed information was collected regarding the activities developed in each building, considering the observable usages: (a) at ground level; (b) on the first floor (or the storey above the ground floor at street level); and (c) the use that prevail (on average) on the storeys above the first floor.

Uses were classified into categories according to their capacity to generate certain movement patterns at different times and days: Commercial (offices, companies), Institutional (government offices and agencies), Shop (retail, commerce), Private service (workshops, barbershops, hairstylists, etc), Community service (religious services, community associations, unions, etc), Health service, Education (schools, colleges, language courses, technical courses, etc), Catering (where the main activity is the sale and consumption of food), Hotel, Leisure and entertainment (theatres, cinemas, bars, art galleries, cybercafés, etc), Transport (stations, harbours, airports, etc.), Industry/handicraft, and Residential. In addition, buildings with little or no use were recorded and classified in the following categories: Under construction, Closed/without apparent use, Utilized as parking lots/garages, Storage and Empty plot. Uses not compatible with this list or unidentifiable from field observation were recorded as Others.

Thematic maps represent the use recorded for each of the edifications (figure 3). These can be re-worked to convey information that may be more or less specific.

Thematic maps were also produced in order to represent the functional categories potentially able to attract associated levels of activities, (Figure 4) maintaining the designations for the following items: Commercial, Private service, Public service, and amalgamating categories which define usages with low affluence of people in relation to the other categories: Storage, Under construction and Parking lots/garages and Closed/without use.

The thematic maps in figure 4 shows a concentration of buildings with a low affluence of users along the blocks closer to the river, where most older buildings of strong historic-cultural appeal are located. It also shows that the distribution of buildings which provide services to the general public crosses over the whole district, tending to concentrate on the north and on the south-centre borders, and alongside Duque de Caxias Avenue, the main neighbourhood artery, which connects the two clusters. Commercial and private services also intensify on this artery radiating from it round all the central blocks. Residential buildings concentrate on the southeastern and northern borders of the neighbourhood.

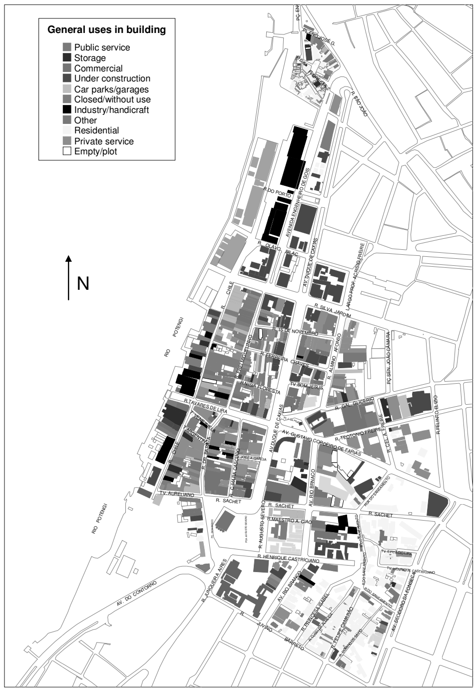


Figure 3: General use on the ground floor

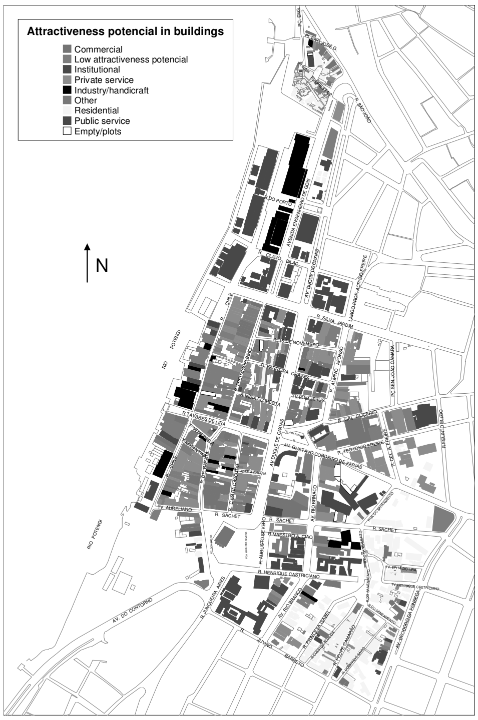


Figure 4: Uses according to the attraction potential in the ground floor

Figure 5 shows some of the specific uses according to the potential of attraction on the floor immediately above the ground floor and the predominant use on the floors above the first floor, where hardly any other use remain but institutional activities, public services and very few residences.

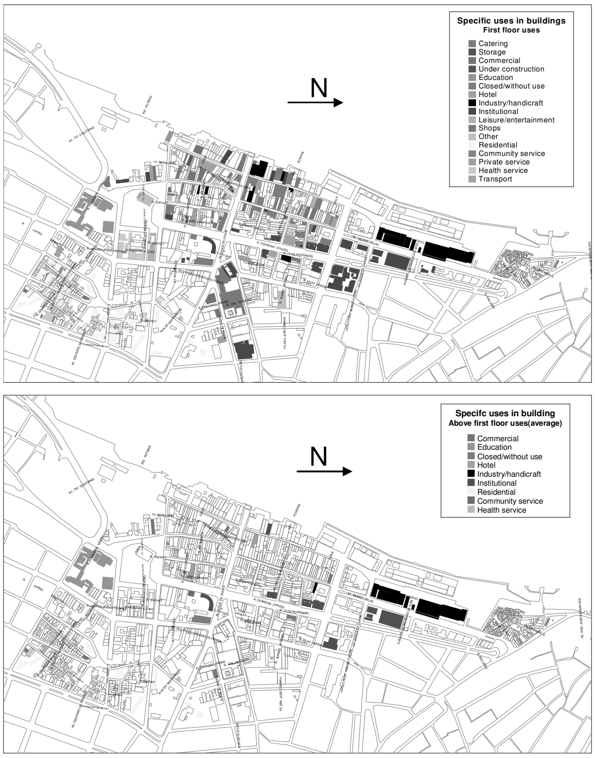


Figure 5: Specific use on the floor immediately above the ground floor and predominant use on the floor above the first floor

2.3 Housing, Diversity and Identity

Identifying qualities inherent to buildings and groups of buildings recommended for preservation

The real estate database resulting from the case-to-case building survey in Ribeira demonstrates that physical, symbolic and functional aspects unite, overlap and can contribute directly or indirectly to the aims expressed in the PRAC/Ribeira's, guidelines:

> "*The aim of the Federal Government is to promote the use and democratic occupation of the central urban areas within Metropolitan Areas, enabling the permanence of the resident population and the attraction of a non-resident population by way of integrated actions to promote and sustain the functional and social diversity, the cultural identity and the economic vitality of these areas*. (Project Guidelines, 2005:2). (Italics by the authors).

Previous studies, (Monteiro, Trigueiro, Roazzi et al, 2003; Elali, 2004) have pointed out that the most evident positive quality recognized for Ribeira is its historic-cultural potential, mainly expressed by the existing built environment, which is symbolic of the hegemonic economic and social position that the district enjoyed in the first half of the 20th century as the active core (or central business district) of Natal.

The acknowledgement of values – such as antiquity and monumentality – of this ensemble is, therefore, the most evident factor of what can be perceived as the cultural identity of Ribeira, one of the central axes of the PRAC/Ribeira. The architectural legacy is also a strong vector for the promotion of sustainable functional and social diversity, as it already accommodate diverse uses – especially commerce and public and private services – being also able to accommodate, in a broad and multifaceted scale, uses considered to be very desirable in rehabilitation schemes, such as housing.

In answer to those propositions, a collection of buildings were recommended for protective actions. They were classified according to formal variables which better encompass potential or existing requirements necessary to meet the above presented aims, from a point of view of the built environment.

These are buildings that either individually or as an ensemble:

> 1. have a strong symbolic appeal, be that potential or acknowledged, and/or contribute to a notion of cultural identity for Ribeira and surroundings – mostly (a) buildings presenting inherent historic value; and (b) groups of buildings which when associated, define a characteristic ambience of a certain location;
>
> 2. present formal and functional attributes that are adequate to house (or carry on housing) uses which are considered important for the generation /maintenance of functional and social diversity and, therefore, of satisfactory levels of urban vitality in Ribeira.

Such buildings were classified in categories, either according to their individual edificial qualities or for contributing, as an ensemble, for the composition of certain sceneries containing a strong identity appeal.

The criteria defining each category, to be explained and illustrated by examples given below, point to certain valuation potentials that demand different levels of attention and protection. The location of the built ensemble, resulting from the identification of categories for buildings recommended as objects of attention and protection, served as basis for redefining the limits of the Ribeira Historic Site.

Defining categories for buildings recommended for preservation, according to their inherent or potential qualities

Approximately 220 buildings were selected as deserving special attention concerning education actions, legal measures and/or public incentives for restoration, preservation or conservation. Categories were defined according to their individual qualities or as part of an ensemble considered for its historic-cultural reference.

> 1. Listed buildings – those that are already object of institutional protection
>
> 2. Antiquity – the built ensemble gathering the largest number of the oldest buildings in Ribeira

3. Landmarks – buildings whose original features are more or less preserved and/or which are recognized as reference points in Ribeira;

4. Rarity – a rare example of a certain epoch in terms of the remaining built stock still existing in Ribeira or in Natal;

5. Successive stylistic references – part of a group of buildings representing successive epochs;

6. Current non-residential adequate use – edifices that accommodate uses that are compatible with the material and symbolic nature of the building and the area;

7. Current residential adequate use – residences that are compatible with the material and symbolic nature of the building and the area;

8. Mixed-use potential – buildings whose exterior appearance and location looks appropriate for mixed use.

2.4 Interpretations and Actions Originated from the Database.

The oldest architectural vestiges of occupation in Ribeira date from the colonial architectural heritage, which goes beyond the date of the Brazilian Independence, (7/ Sept/ 1822), coexisting until the second half of the 19th century, alongside the then prevalent eclectic tendencies, particularly the classicizing ones. Rua Chile concentrates the greatest majority of architectural vestiges showing colonial and neo-classic heritages, which survived in the neighbourhood or even in Natal

Therefore, institutional protection is recommended for the whole old ensemble of Rua Chile, (Figure 6) where the largest number of buildings keep morphological attributes representing the most ancient occupation epoch in Ribeira.

Figure 6: Buildings that present stylistic attributes inherited from colonial times (Rua Chile)

Environmental studies demonstrate that some buildings are perceived and mentioned as urban landmarks, in symbolic terms, as a historic cultural reference or as demarcators of urban intelligibility. Such buildings, here considered as urban landmarks, are concentrated round squares and along the main thoroughfares (Figure 7).

Figure 7: A landmark in Ribeira (Alberto Maranhão Theater)

It is proposed that landmark buildings now belonging to government agencies or which house public institutions be listed for preservation while the remaining others are recommended for examination pending institutional protection.

At different points of Ribeira there are buildings whose formal aspects belong to stylistic tendencies that have almost completely disappeared in Natal The ones located close to listed buildings and landmarks are recommended for preservation listing. The remaining ones are recommended for incentives to encourage the maintenance of their exterior built shells.

The neighbourhood also contains an expressive number of non-residential buildings whose formal vestiges characterize successive epochs of the Ribeira occupation, demarcating stages of such occupation and contributing to the identification of places of touristic cultural interest, even if individually considered not all those buildings present notable architectonic qualities.

Some buildings that are well conserved and well preserved in their original formal attributes accommodate uses considered adequate for their physical structures and for the mixed use pattern that is being devised for Ribeira. Incentives are recommended for the conservation of these buildings' physical attributes and function.

Dwelling houses dating from successive periods of time located in the highest area of Ribeira (where middle class housing predominate) but also on the river bank, as part of one of the oldest slum communities of Natal – the Maruin – deserve attention because they combine some key aspects proposed in the PRAC – Ribeira: the maintenance of residential *mixité* and the possibility to attract new residents to Ribeira.

Moreover, in the referred residential enclaves, the private spaces of the houses and the public spaces of the streets though well demarcated, interface directly, thus contributing to generate spatial properties of co-presence and visibility, which are deemed favorable for fostering urban vitality. Such properties have often been sought after in urban redesigns of derelict areas (not always successfully) as they may induce uses that benefit from the proximity of homes (local commerce/service outlets), foster animation by maximizing random encounter opportunities among dwellers and between these and outsiders, and reduce opportunities for antisocial behaviour by enhancing surveillance concerning the presence and movement of strangers by residents.

Attention is, therefore, recommended, regarding the possibility of establishing incentives for the maintenance of residential use, built shells, land parceling, building/plot relationship and private/public interface in these areas.
Buildings dating from successive periods of time have built shells that appear compatible with residential or mixed use (commercial/service on the ground floors with residences above). The ones located on several streets of Ribeira where the re-introduction of residents is desirable are recommended for inspection of their physical structures with the aim of verifying the possibility of adapting their upper floors for residential use.

Housing in the old town centre

Unoccupied or under-occupied buildings were selected for residential use in tune with the project's guidelines for enhancing residential use in the area, as part of a strategy for its rehabilitation They were classified as: (a) buildings recommended for conservation; and (b) buildings that may be converted, reconstructed or substituted.

3. SERIDÓ'S BUILT HERITAGE

3.1 About the Region of Seridó and its Architectural Legacy

The region of Seridó (Figure 1) concentrates some of the oldest occupation nuclei in the state of Rio Grande do Norte. An important stock of buildings dating from successive stages of occupation can still be found in the towns' old centres although in progressive and rapid pace of disappearance, since the 1970s. This legacy, added to the fact that the region has always been a focus of socio-cultural dissemination, grant its remaining stock of older buildings the status of acknowledgement and preservation deserving artefacts.

Historically linked to cattle breeding, these towns changed significantly along the 19th century, accompanying the growth of commerce and the development of cotton wool plantations in Rio Grande do Norte, thereafter, functioning mainly as supporting centres for farming activities until well into the 60s. From then on, in the trail of the urbanisation impetus that swept through Brazil, and of the local decline of beef and cotton production, both unable to compete with that of other origins, those towns rapidly turned into local centres for tertiary activities. Former residential buildings in the original town cores, where 19th and early 20th century remains can be found are being increasingly converted into shops and service outlets. As a general belief that the look of domestic buildings is incompatible with commercial use predominates in Brazil, original façades are giving way to wide glass-paned doorways and windows, topped by commercial ads, and built shells are being hollowed.

In addition to the loss of decades of socio-cultural practices crystallised in spatial structures and building materials, this process seems particularly alarming in view of the fact that, in this country, although architectural integrity has not yet fully become a material asset, architectural mutilation is a strong factor for devaluation and urban decay.

If the assumption advocated by various authors that the world is seeing an almost religious worship of architectural heritage is to be believed, such practices place the majority of the Brazilian population outside the world. On the other hand the increasing number of courses/events dedicated to discussing heritage and of conservation plans being produced suggest that the gap between the way scholars/technicians and the general public perceive the built environment is widening.

Riegl (1984) responds to the attitude of attributing superiority to all things new at the expense of older ones. The quality of newness can be easily appreciated by anyone despite his or her education degree or specific training. In Brazil this is often manifested in pointless modification or destruction of old buildings solely to update facades to the newest trend .this, added to other factors owing to contemporary needs or tendencies (i.e. car ports, growing preferences for apartment living, the emergence of new neighbourhoods that tend to attract more public investment) combine to foster a grim picture for older town centres.

In the face of such prospect an attempt to disseminate knowledge about the built heritage of Seridó by "translating" architectural information, normally confined in Brazil within the academic sphere, to the larger community of internet users was developed by researchers engaged in heritage studies at the research group on morphology and use of architecture – MUsA/UFRN (Figure 8).

Figure 8: Initial page of the site about the remaining built heritage at the end of the 20th century in older towns of Seridó (http://www.musa.ct.ufrn.br/bdc/)

3.2 Inventário de uma Herança Ameaçada (An Inventory of a Threatened Inheritance)

The research and outreach project entitled "An inventory of a threatened inheritance" (Inventário de uma herança ameaçada) comprised an extensive inventory developed from 1997 through 2001 (Trigueiro et al) in 22 towns of Seridó, aiming at: (1) registering pre-modernist and modernist buildings still surviving at the end of the 20th century; (2) identifying their period of construction; and (3) ascertaining the state of preservation of their original formal features and conservation of their physical factories. The information resulting from the inventory is presented in the site www.musa.ct.ufrn.br/bdc.

The site offers a panorama of the pre-modernist and modernist architectural ensemble dating from the late 18th through to the mid 20th century still surviving at the end of the 20th century in Seridó. It is, therefore, an extensive inventory whose merit is essentially that of recognizing a legacy in extinction at a moment considered crucial as it may be the last opportunity to visualise that patrimony in the living urban scene. By offering simple tools for identifying architectural vestiges that predominated in successive time periods – colonial, eclectic and modernist – from the form of their represented built shells we seek to contribute towards narrowing that gap in the hope that the images in the web may contribute to enhance awareness concerning the fading architectural heritage of Seridó, especially among young internet users.

The database concerning Caicó – the oldest and most important town in the region – was built between 1996 and 1998. At the time of the inventory, the studied towns were distributed into "homogeneous zones" defined in official government data, according to proximity, geographical similarities and the area of influence of some larger towns, as follows:

- Zona Homogênea de Caicó: Caicó, Cruzeta, Acari, Ipueira, Jardim de Piranhas, Jardim do Seridó, Ouro Branco, Santana do Seridó, São Fernando, São José do Seridó, São João do Sabugi, Serra Negra and Timbaúba dos Batistas;
- Zona Homogênea de Currais Novos: Currais Novos, Carnaúba dos Dantas, Parelhas and Equador;
- Zona Homogênea das Serras Centrais (central sierras): Jucurutu, Florânea, São Vicente, Lagoa Nova and Cerro Corá.

Proceedings applied in the field work included: (1) the identification of each building in maps that displayed the areas occupied until the 60s in the surveyed towns; (2) photograph records of the buildings (individually or grouped in rows) which conserved some or all original formal features visible from the street; and (3) synoptic file records containing address, location sketch, date, number of negative exposure, photographer, key formal features, state of preservation of original features, the state of conservation of built fabrics, and observation notes.

The data gathered in the field journeys was organised into: (1) a film negative file; (2) a photograph file; (3) a general information paper file catalogue that combines image, identification records and formal features; (4) a map file; and (4) a reference table. The analysis of the data recorded allowed for an estimated identification of the buildings into successive time periods defined as: (1) colonial – when formal features inherited (though not necessarily built) from colonial times predominate; (2) eclectic – when formal elements pertaining to or inspired by any neo-tendencies, from neo-classical to neo-colonial, as well as Art-nouveau and Art-deco, individually or combined, predominate; and (3) modernist – when either the simple geometrical nudity of modernist or the array of elements associated nationally and/or regionally with the dissemination of modernist (or functional) taste predominate in the angles of the buildings visible from the exterior.

Most of the above mentioned information is being progressively transferred to a digital bank created for the purpose so that the data can be easily exported for formats compatible with various soft wares that allow for GIS mapping and internet communication. Some of these procedures are now concluded, others being still in progress as, for instance, the map files whose sketches utilised in the field work are being transferred to a GIS-supported database. The bank is the basis of the inventory website, which can be provisionally accessed at the address www.musa.ct.ufrn.br/bdc.

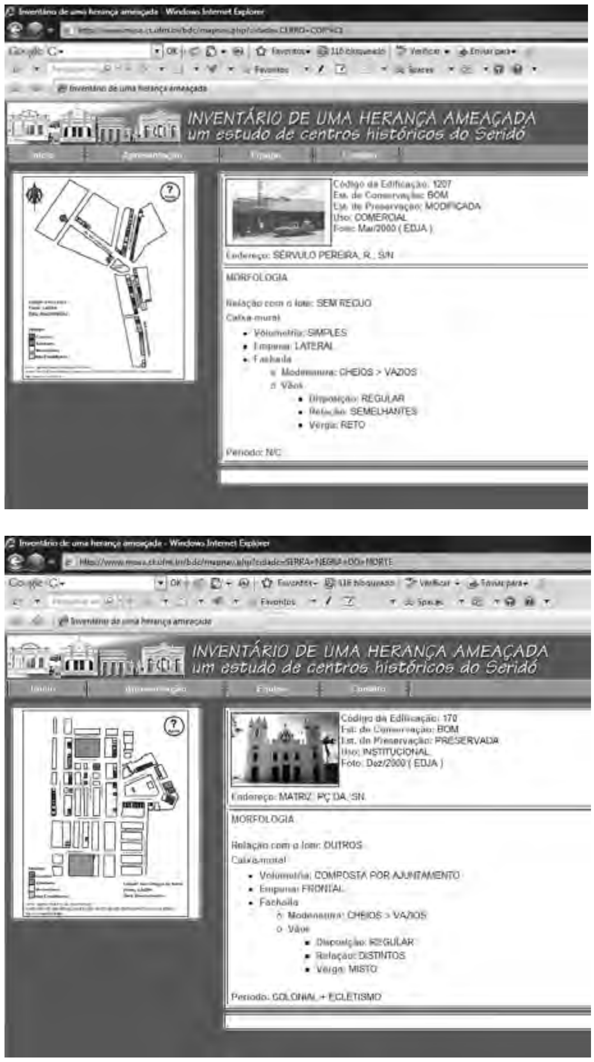


Figure 9: Town maps and identification file of a building in Cerro Corá (above) and Serra Negra (below)

The site presents interactive town maps (Figure 9) with each building colour-coded according to the stylistic trend expressed in their exterior built shells that conveys information about its time of construction (or major stylistic update). By clicking on each building its identification file (location, formal features etc) is displayed. It shows the buildings images and other aspects such as state of preservation of original features, and state of conservation of built fabrics. Navigation is possible through the addresses, the map recordings and some key-words. Printing is also possible.

Although still under construction the site has been frequently used to provide data for studies of diverse environmental-related natures, directly or indirectly associated to the effects that the formation and transformation of the urban occupation may have over the construction, modification and substitution of the built environment. Architecture, History, Geography, Tourism and Law students – at undergraduate and graduate levels – have reported utilisation, or enquired about information existing (or lacking) in the site.

4. TOWARDS AN ATLAS OF ARCHITECTURAL HERITAGE IN RIO GRANDE DO NORTE

4.1 Future Actions

The database containing information about the buildings identified in the research projects *Plano de Reabilitação de Áreas Urbanas Centrais – PRAC-Ribeira* and *Inventário de uma herança ameaçada* result from academic activities of various natures, which include teaching, researching, outreach and consultancy, conducted by professors, researchers and students, working in collaboration at MUsA-UFRN. Most of this material had long been exchanged and utilised by fellow academics in hard and digital formats but used to remain largely restricted to members of our university, and especially of our department. We think it is high time that this knowledge be shared with all those interested in the built environment by taking advantage of current ICT technologies, which have become fairly well accessible even in the hinterland towns of Brazil.

It is here hoped that free access to this information may also facilitate the construction of other sources of knowledge dissemination about our built heritage. Having this in view we are beginning to develop the basis of what is expected to become an Atlas of Architectural Heritage in Rio Grande do Norte to be displayed on the web. This databank shall complement the already existing atlas of socio-economic aspects in the Northeast of Brazil, constructed by researchers participating in a joint project that involved the INPE (National Agency for Space Research) and the UFRN.

We seek to build the atlas in a chronological order of urban occupation, starting with the original old centre of Natal (first urban settlement in the state, founded in 1599), proceeding to other neighbourhoods in this town (i.e. Ribeira), and then extending the panorama to include other old towns of Rio Grande do Norte. If no efficacious actions come out of this effort in terms of a more respectful attitude towards whatever remains of our depleted built patrimony, the notion of having not withheld the product of decades of academic laborious work may serve as consolation.

REFERENCES

Elali, G., 2004. Imagem sócio-ambiental de áreas urbana: um estudo na Ribeira, Natal, RN-Brasil. *Revista Psicologia para América Latina*, número temático La Psicoligá y la Ciudad.

Monteiro, C.; Trigueiro, E.; Roazzi, A. et al, (2003). *Morar no centro: pesquisa de demanda habitacional no centro histórico de Natal*. Research Report. Natal: UFPE/UFRN.

Project Guidelines: Plano de reabilitação de áreas urbanas centrais: Ribeira, (2005). Natal: Prefeitura Municipal de Natal.

Riegl, A., *Le culte moderne des monuments. Son essence et sa génèse*. Éditions du Seuil, Paris, 1984.

SPATIAL PATTERNING OF EPIPALAEOLITHIC-EARLY NEOLITHIC SITE STRUCTURE OF IZEH PLAIN, SOUTHWESTERN IRAN

Niknami, Kamal Aldin[a] and Jayez, Mozhgan[b]

Department of Archaeology
University of Tehran
Enghlab Street
Tehran, Iran
[a]E-mail: kniknami@ut.ac.ir
[b]E-mail: Mozhgan_78224@yahoo.com

KEYWORDS: Epipalaeolithic, Early Neolithic, Spatial analysis, Settlement pattern, Principal component analysis, Regression analysis

ABSTRACT:

The aim of this work is to show how the most advanced technology together with spatial analysis can be usefully employed to investigate archaeological phenomena. This paper presents a summary of the basic methodology and some results of a recent (2006-2007) archaeological project concerning the search for Epipalaeolithic/Early Neolithic site distributions of Izeh plain. The primary goal of this project was to detect Epipalaeolithic/Early Neolithic site locations and to provide the necessary spatial models to understand the distribution pattern and structure of sites. We applied several analytic and search procedures to a well documented Epipalaeolithic/Early Neolithic case in order to infer prehistoric behavior in the organization of space. we detailed the results of an intensive ground survey of Epipalaeolithic/Early Neolithic sites designed to determine: 1) the spatial pattern of sites at a landscape scale; 2) whether a positive association exists between the density of sites and environmental variables and if so, at what scale this relationship is strongest; 3) to develop a computer based archaeological site management system.

1. INTRODUCTION

This paper is an analysis of site structure as it applies to social and behavioral organization within Paleolithic hunter-gather camp sites. A complete understanding of site structure requires the examination of high-definition sites that are very well preserved in terms of their spatial integrity. The spatial structures of Izeh plain sites are comparable to the spatial structures of many other Paleolithic sites and certain types of ethnographic hunter gatherer sites that can be studied for a spatial examination of site structure. It can be argued from the potentials of sites that these sites are optimal resources for spatial analysis but there has always been the problem of choosing the most effective techniques in performing the spatial analysis. New techniques are continuously being developed, each of which add to the ability of accurately analyzing spatial data, but the most important advance in the recent past has been the development of new tools that can be used in spatial analysis. Geographic Information Systems are particularly well suited for the spatial analysis of Paleolithic sites since they are able to quickly and easily store and work with the large amount of spatial data that results from the survey of Paleolithic sites New techniques of spatial analysis available through Geographic Information Systems allow for a much better understanding of spatial relationships then was previously possible. Once the spatial data from these sites is analyzed another problem arises of how best to interpret the results. The results may be interpreted based on a variety of assumptions about the processes that produced the spatial pattern. A particular model of site structure may also order an interpretation of spatial data from hunter-gatherer sites. Finally, the assumptions made about Paleolithic hunter-gatherer camp sites as well the site structure models applied to them have changed dramatically over the past thirty years due to more and more advanced archaeological research and the incorporation of ethnoarchaeological observations (Keeler 2003). The analysis that we perform on the data from Izeh plain Epipalaeolithc-Early Neolithic will utilize some of the most advanced techniques of spatial analysis available in Geographic Information Systems. The results of this analysis will then be interpreted using the most current models and assumptions of such site structure in order to better understand Epipalaeolithic behavioral and social organization at this region.

2. RESEARCH BACKGROUND

In 2007, a survey was done in Izeh Plain (Northeastern Khuzistan, Iran) in order to complete our knowledge about Stone Age sites of the area and analyze their spatial pattern of distribution. The majority of the stone artifacts collected during the survey show an Epipalaeolithic/Early Neolithic industry.

In the Near East there are two major regions in which Epipalaeolithic studies are extended: 1. Levant (see Bar-Yosef, 1998; Henry, 1989; Belfer-Cohen, 1991; Goring-Morris, 1987), 2. Zagros Mountains. In the latter, the typical Epipalaeolithic industry is named "Zarzi" based on a cave with the same name which was excavated by D. A. E. Garrod (Garrod, 1930). Zarzian Industry is found not only in sites which are located in Central Zagros (Olszewski, 1993a; Wahida, 1999; Young & Smith, 1966; Braidwood & Howe, 1960), but also in some sites in southern Zagros (Smith, 1986: 30; Tsunki et al., 2007; Dashtizadeh, 2003).

Izeh plain was first surveyed by a joint team of archaeologists conducted by Henry T. Wright in 1970s in a salvation project (Wright, 1979). During that project 18 Epipalaeolithic sites were discovered. Wright mentioned in his report that he had surveyed only 30% of the foothills; in consequence he was not able to find any kind of "settlement pattern" for the Epipalaeolithic sites. It is noteworthy to mention that Epipalaeolithic sites are found not only in Izeh plain, but also in some areas around it, like the area which was surveyed during salvation project of dam Karun 3 and is located in southeastern Izeh and in Dashte Gol which was surveyed as a part of salvation project by Wright (1979).

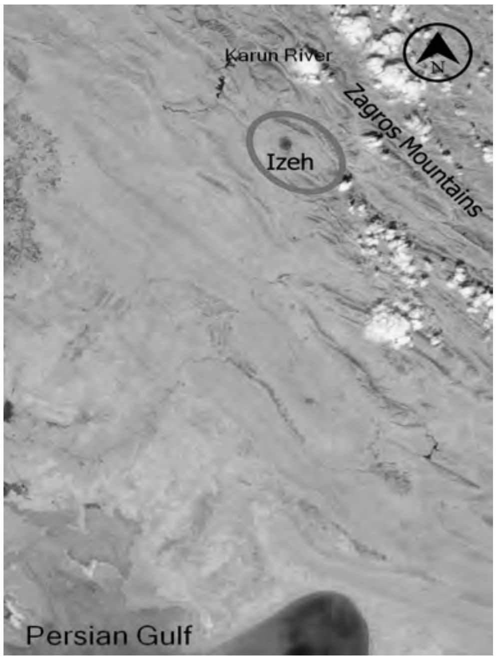


Fig. 1. Satellite Image of Southwestern Iran

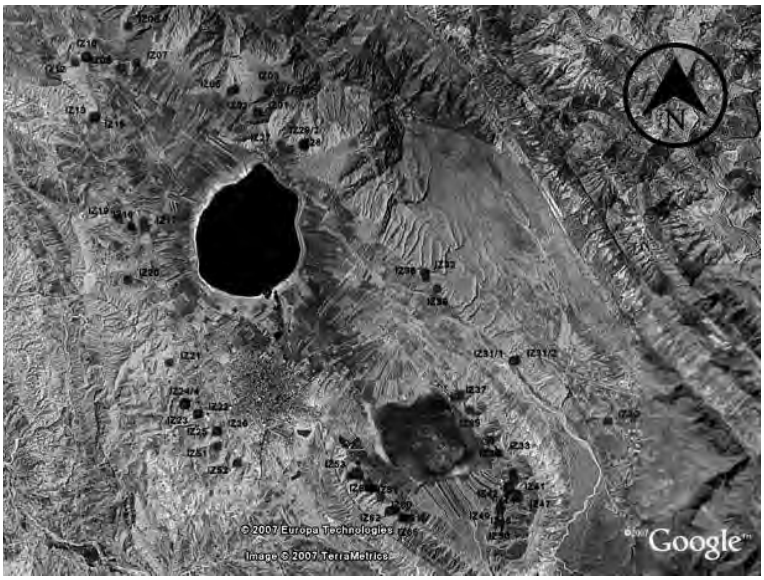


Fig. 2. The location of Stone Age sites in Izeh plain discovered during 2007 survey project

3. MATERIAL AND METHOD

Izeh plain is an alluvial plain located between N 31 54 – 31 47 and E 50 00 – 49 47 in Northeastern of Khuzistan province, in Iran (fig. 1). It is averagely 750m above sea level and extended in 140 km2 (23 km NW to SE, 12 km N to S). It is located in a semi-steepy region and generally has the lowest temperature in the Khuzistan. It is also internally drained and two lakes named "Miangaran" and "Bondun", first in the central part of

Izeh and the second in southeastern part, play a key role in ecological characteristics of the plain.
In 2007, about 125 km2 of Izeh plain was surveyed between maximum elevation of 1159 and minimum of 854 m above sea level which contains lowland skirts, foothills, and those parts of plain which are rather intact from floods, industrial constructions and husbandry. As a consequence 54 sites were discovered (fig. 2) comprising Epipalaeolithic/Early Neolithic cultural periods, based on typology of the stone artifacts (fig. 3, 4). Some of the sites were previously discovered by Wright (such as IZ05, IZ20, IZ24, IZ34, IZ35), but many of them are recorded for the first time.

Among the sites are 6 groups of cave and eshkaft[1], 5 groups of eshkaft, 1 group of cave and eshkaft and rock shelter, 27 eshkafts, 3 caves, 7 rock shelters, 2 open air sites and 1 rocky hill[2].

Fig. 3. Stone tools from Izeh Plain

The stone Industry of the sites of Izeh is mostly based on production of bladelets reduced from different bladelet cores. There are different tools like small scrapers, thumbnail scrapers, small borers, small burins[3], backed blades and bladelets, few geometric microliths in the assemblage, but the majority of the tools are retouched bladelets. There are also different bladelet cores like single/double platform cores and conical microblade cores (fig. 3, 4).

Fig. 4. Stone tools from Izeh Plain

It seems that the whole assemblage shows an Epipalaeolithic Industry (although the geometric microliths are rare) and Early Neolithic (according to few bullet cores and many bladelets).

4. STATISTICAL ANALYSIS

We used Analytical method of PCA (principal component analysis) for reduction of our data and realizing the most important factors in typological and technological distribution of stone artifacts. PCA is designed to reduce the number of variables that need to be considered to a small number of indices that are linear combinations of the original variables. PCA provides an objective way of finding indices of types so that the variation in the data can be accounted for as concisely as possible. It may well turn out that two or three principal components provide a good summary of all the original variables. At first qualitative characteristics of the artifacts have been coded and identified in ratio scales. The method explains the role of some of factors which can describe some variants in relation to other variants. By applying varimax rotation method, major factors were found, from which factors 1, 2 and 3 are explained here (fig. 5, 6, 7).

First factor shows some characteristics which represent mostly flake production and utilization. This is confirmed by the presence of flake cores, mixed cores, corticated core flakes and with most blanks being flakes. The presence of some retouched blanks (flakes, Proximal/ distal parts, flake blades, bladelets) demonstrates the utilization of productions in situ. Actually, the assemblage shows the process of reduction of flakes from cores, retouching and using them. Different flake cores (mixed irregular core, double platform flake core, single platform flake core, single platform mixed core) demonstrates that the

[1] "Eshkaft" is a shallow cave; the name is given from local language of the Baxtiary nomads of Izeh.
[2] For full details of the sites and their stone artifacts see Jayez, 2008.
[3] They are burins which are small in size. No "microburin" was found in the assemblage.

production of flakes is not steady and the presence of corticated proximal/distal parts in the sites show that the implements had gone under further work in the sites. Corticated retouched bladelets in the assemblages hint at the point that the sites were not only a workshop of flake production, but a place to use and retouch other implements. This factor is mostly prominent in site 36 (IZ42/43/44) which is located in the marginal lands of Lake Bondun in south eastern of Izeh plain, but there are also some sites in north western end of the plain (IZ07, IZ09,10,11, IZ12) with a distance of 3 km from Lake Miangaran and few ones in south western of the plain in the vicinity of the city of Izeh (IZ21). The rest of sites are all located in the strait of Bondun Lake.

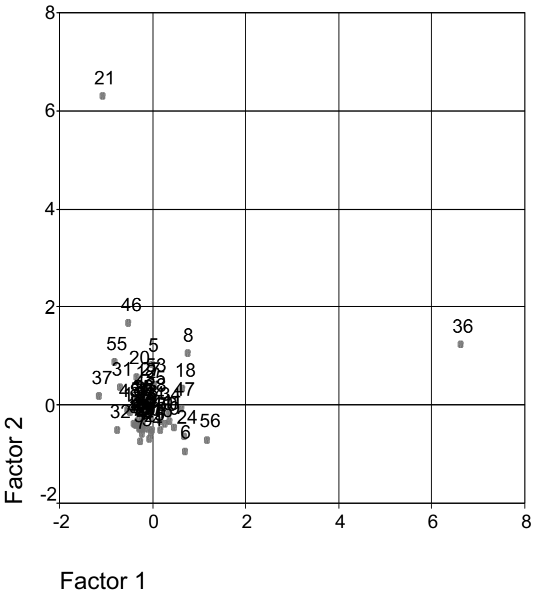


Fig.5. Plot of the first two principal axes from a PCA analysis

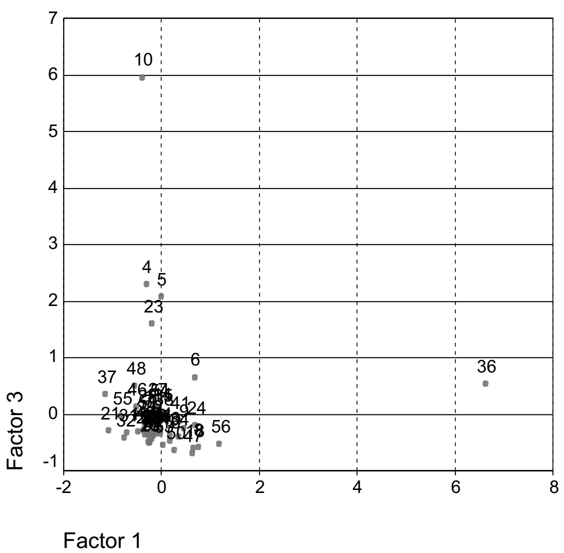


Fig.6. Plot of the first and third principal axes from a PCA analysis

Second factor shows the sites which by no mean represent the process of reduction of blanks or tool making, especially cores are rare in this group. Blanks show a high variation (bladelets, blades, proximal/distal parts of bladelets, etc.) and accordingly different activities having taken place in the sites. Debris increases the probability of reduction on those implements which got blunt or unusable. These blanks and tools were surely made in other places (first or third group sites) and transmitted to be used in other places (second group). This factor is mostly prominent in site 21 (IZ27), but other sites are distributed almost equally in northwest of the plain (IZ06), north (IZ05), east (IZ32, IZ38), south west (IZ24, IZ26) and again in the strait of Lake Bondun which shows a high density of the sites of this group. The only part of the plain which lacks the sites of this group is eastern part of the plain.

Third factor shows clearly the sites which were workshops of blade/bladelet production according to different bladelet cores (single/double platform cores, conical microblade cores, bipolar cores, mixed cores and irregular cores). On the other hand there are few blade/bladelets in the assemblages. The small number of blade/bladelets in this group shows that they were transmitted to other places (second group/ first group or even out of the plain) after production, although some of them were utilized in situ and this is confirmed by the presence of few microliths. This factor is mostly prominent in site 12 (IZ17) in western margins of Lake Miangaran. The site is a rocky hill which at the moment is a religious shrine. Other sites of this group show a low density and are located along northwest-southeastern foothills (IZ12, IZ13,14,15, IZ24) and a site in north of Lake Miangaran (IZ27). Four other sites of this group are located inside of the strait of Bondun Lake.

It is noteworthy to mention the variation of stone artifacts density representing in the site. The regional analyses seem to indicate that the geomorphology and ecological settings of the region would have played a large role in the site-placement process, along with an increased dependence on access to raw materials. As site became larger and more permanently settled, we would expect an increase in the role played by habitability factors namely environmental suitability. A multiple regression analysis was carried out to determine if the data were suitable for creating accurate values (statistically derived weighting values) and, therefore, an inferential probabilities model. The values for each independent variable at each site were calculated based on the assumption that one of independent variables (sea level elevation, local elevation, proximity to water resources and slope) may potentially affected the density of tools (dependent variable), exposed on the surface of sites. Though there was a total of 54 sites, the variation in the sample was such that in all samples (except proximity to water sources) less than %27.5 of the variation in dependent variable could be explained with a multiple regression analysis. The adjusted r^2 value could be increased by dropping independent variables, but it appears that just the density of tools can be explained by straight linear relationship to the water proximity variable. The other variables by no mean can explain easily the variant characteristics of independent variable. Likely this is due to the homogeneous nature of the environment of the study area, irregular survey coverage, the inability to accurately model previous survey areas and the inconsistency in assuming that the site size or site density accurately represent intensity of artifacts. Applying a wide range of variables with controlled weights will allow us to directly access known or hypothesized settlement strategies with the archaeological data.

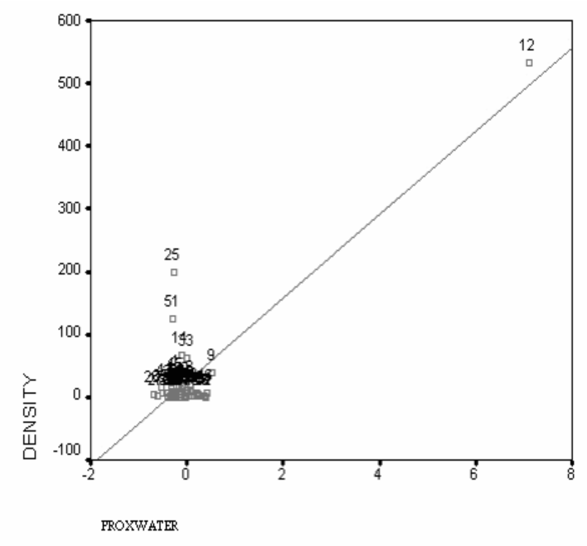


Fig.7. Straight linear relationship between density of artifacts and distance to water sources

5. DISCUSSION

Izeh plain shows a homogenous geomorphologic and ecologic landscape which could be the reason for almost homogenous distribution of Stone Age sites.

Actually two seasonal lakes in the plain and geological characteristics of the area which has led to formation of many caves, eshkafts and rock shelters (many of which are still in use as fold) provided a potentiality for hunter-gatherers in transition to extended ways of food-collecting and food-producing.

The picture of functional distribution of sites which was reached from statistical analysis in this study shows that different functional sites (described above) are distributed randomly and with few distances from each other, and that transition between these sites shows no clear pattern, except that the majority of sites with various functions are located in the strait of Lake Bondun.

Spatial distribution of sites in Izeh plain shows a homogenous accessibility to raw material for producing stone artifacts. This issue will be analyzed in the future studies. It seems that the most of hunter-gatherers of Izeh had been semi-sedentarists according to size and relational distribution of different types of sites which are mostly caves and eshkafts.

We could put in a model concerning settlement structures and various characteristics of the observable activities in the region. Various models are presented about description, prospect of the organization, composition of stone implements and their relation with basic economy of hunter-gatherer bands. They started with Binford (1979) and Renfrew (1977) in which they explained the relation between transmission patterns with accessibility to raw material. Today these models have evolved to explain with more complexity about relations between individual and band movements of hunter-gatherers, their environment and required recourses in their subsistence (e.g. see Jones et al., 2003; Brantinghom, 2003, 2006; Shott, 1994). Generally in these models different activities of hunter-gatherers is assessed based on the composition and structure of raw material and stone implements and from various characteristics of the implements like size analysis, typology and their function. In most of these models recognition of sedentary or semi-sedentary groups is through the factors and the pattern of their accessibility to raw material. One of their characteristics is that they discard overused old implements and replace them with new ones, and that is because of their access to different raw material resources due to their mobility. Sedentary societies, unlikely use both poor raw material in their access and high quality material which they gain through trade in sometimes remote distances which leads to formation of non regional industries but with higher quality compared with the local one (Binford, 1980). Totally, stone assemblages from settled societies show more variability in raw material from different distant sources. Another approach in distinguishing semi-sedentary societies is that the average size of flakes changes parallel to increase of distant, it means that flakes occasionally decrease in size due to continual retouching.

Spatial distribution of sites in Izeh Plain which has an almost homogenous landscape shows that the people in the plain had an almost equally access to raw material. It is not still clear if all of the implements are from local recourses or they are imported from other distances, but the implement show a high resemblance to local recourses which supports from the local recourse utility hypothesis. Unfortunately the exact subsistence of Epipalaeolithic people of the region has not yet been demonstrated due to lack of excavations and we can not speak of nutritional recourses with certainty. Although such studies need other field works to be done, but it is clear that the environment of Izeh Plain has a great potentiality in providing various foodstuffs.

It seems that the majority of sites in Izeh Plain show semi-sedentary people, this is supported by small size of the sites that are probably due to continuous movements in the way of which the sites are formed arbitrarily in the Plain (note that there have been found few sites in central parts of the plain because these parts have been under various exploitation for long times, i. e. agriculture, husbandry or rural settlements with great erosion). Also most of the sites are caves or "eshkafts", open air sites are rare which stops us from concerning the people in Izeh Plain as full mobile bands. Although stone implement analysis discussed in this study emphasizes semi-sedentary people in Izeh Plain, we should bear in mind that various functions of sites shows some kind of exchange relations between sites in a way that raw material in those sites with stone implement producing characteristics were reduced to implements which were transmitted to other sites in order to be utilized in different ways.

6. CONCLUSION

The spatial model presented here enables a set of archaeological data to be meaningfully interpreted with respect to human behavior during the Epipalaeolithic-Neolithic of southwestern Iran. The model is rather an interpretation of the processes by which the archaeological record was formed in the Izeh plain. In undertaking this task the model addressed a problematic area for spatial modeling: the activities of semi sedentary societies. The outlined applications of the model are preliminary, but suggest the potential long-term regional scale data to the interpretation of human behavior and landscape exploitation strategies.

One conclusion is proposed:
Tool making strategies were restricted to specific sites. The human populations were apparently exploiting the full range of available environment during the resource acquisition activities. These activities may correspond with the social support system as reflected by stone tools production technology and exchange patters. To emphasis on spatial modeling procedures in future studies, more survey data as well as chronological and environmental data are needed. This trend will shift attentions from identifying archaeological sites to the processes which reveal those sites.

7. REFERENCES

Bar-Yosef, O., 1998. The Natufian Culture in the Levant, threshold to the origins of agriculture. *Evolutionary Anthropology: Issue, News and Reviews 6*, pp., 159-177.

Belfer-Cohen, A., 1991, The Natufian in the Levant. *Annual Review of Anthropology 20*, pp., 167 – 86.

Binford, Lewis R., 1979. Organization and formation processes: Looking at curated technologies. *Journal of Archaeological Research 35*, pp., 255 – 273.

Binford, Lewis R., 1980. Willows smoke and dogs tails: hunter gatherer settlement systems and archaeological site formation. *American Antiquity* 45, pp., 4 – 20.

Braidwood, R. J. and B. Howe, 1960. *Prehistoric Investigation in Iraqi Kurdistan.* Studies in Ancient Oriental Civilizations 31, Chicago; Illinois: The University of Chicago Press.

Brantinghom, P. J., 2003. A neutral model for stone raw material procurement. *American Antiquity 68*, pp., 487 – 509.

Brantinghom, P. J., 2006. measuring forage mobility *Current Anthropology (47)*, pp., 435–459.

Dashtizadhe, A., 2003. A survey of Neolithic settlements of Kazerun Caves. In S. Zare (ed.), Proceedings of The Second Congress of Young Archaeologists in Iran, Tehran: CHO, pp. 25-48. (in Persian).

Garrod, D.A.E., 1930. *The Palaeolithic of Southern Kurdistan: Excavations in the Caves of Zarzi and Hazar Merd.* Bulletin no.6, American School of Prehistoric Research, New Haven, pp. 9-43.

Goring-Morris, A. N., 1987. *At the Edge: Terminal Pleistocene Hunter-Gatherers in the Negev and Sinai.* Oxford: BAR International Series 361.

Henry, D. O., 1989. *From Foraging to Agriculture: the Levant at the End of the Ice Age.* Philadelphia: University of Pennsylvania Press.

Jayez, M., 2008. *Stone Age Sites of Izeh Plain: Discovery and Spatial Analysis*. Unpublished M.A. Thesis, Department of Archaeology, Faculty of letters and Humanities, University of Tehran. (in Persian).

Jones, G.T., C. Beck, E.E. Jones & R.E. Hughes. 2003. Lithic source use and Palaeoarchaic foraging territories in the Great Basin. *American Antiquity 68*, pp., 58–83.

Keeler, D. 2003. Intrasite spatial analysis of a Late Upper Paleolithic French sites using geographic information systems. *Journal of World Anthropology: Occasional Papers, Volume III, Number 1*, pp., 1-40.

Olszewski, Deborah I., 1993. Zarzian microliths from Warwasi rock shelter, Iran: Scalene Triangles as Arrow Components. In Hunting and Animal Exploitation in the Late Palaeolithic and Mesolithic of Eurasia; edited by G. L. Peterkin, H. M. Bricker and P. Mellars. Archaeological

Papers of American Anthropological Association 4: 199 – 205.

Renfrew, C., 1977. Alternative models for exchange and spatial distribution. In Earl, T. K. and J. E. Ericson (eds.), Exchange System in Prehistory, NY, Academic Press, pp. 71-90.

Shott, M. J., 1994. Size and form in the analysis of flake debris: review and recent approaches. *Journal of Archaeological Method and Theory 1*, pp., 69 – 110.

Smith, Philip E. L., 1986. *Palaeolithich Archaeology in Iran*. The American Institute of Iranian Studies, Monograph No. 1, Philadelphia: The University Museum, University of Pennsylvania.

Wahida, Ghanim, 1999. The Zarzian industry of the Zagros Mountains. In W. Davies, R. Charles, (eds.), Dorothy Garrod and the Progress of the Palelolithic Studies in the Prehistoric Archaeology of the Near East and Europe, Oxford: Oxbow Books, pp. 181-208.

Wright, Henry T. (ed.), 1979. *Archaeological Investigations in Northeastern Xuzestan*, 1976. Research Report in Archaeology; Contribution 5, Technical Report 10, Museum of Anthropology, The University of Michigan, Ann Arbor

Young, T. C. and E. L. Smith, 1966. Research in the prehistory of Central Western Iran. *Science (153)*, pp., 386-391

8. ACKNOWLEDGMENTS

We would like to thank Mr. J. Mehrkian (Moderator of Izeh Cultural Heritage Center), Mr. N. Salahshur and Mr. M. Faraji, for participating in fieldwork, Mr. F. Biglari and Mr. A. Dashtizadeh for helping in stone artifact studies.

Semantics and Standards

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

SHARING 3D ARCHAEOLOGICAL DATA: TOOLS AND SEMANTIC APPROACHES

A. D'Andrea

CISA, Università di Napoli L'Orientale, Vicoletto I, S. Maria ad Agnone 8, 80139 Napoli, Italy - dandrea@unior.it

KEY WORDS: Laserscanner 3D, archaeological documentation, standard, interoperability, 3DPDF, GML

ABSTRACT:

Archaeological documentation is the core process of archeological investigation on the field. Digital evolution has dramatically increased the quality and quantity of archaeological documentation, as far as digital photos and drawings produced with CAD techniques are concerned. Composite computational information systems using different platforms, software and, mainly, conceptual data representations have been implemented to manage this great quantity of information.
The availability of new technologies for 3D data acquisition is modifying the paradigm of 3D objects and consequently the work in the field; in particular, the LaserScanner 3D tool is able to gather and store information on the geometry and features of the scanned objects. In the archaeological domain, researchers are experimenting the application of this new equipment to provide complete documentation of ancient monuments (walls, buildings, caves, etc.). Recently archaeologists have started using LaserScanner 3D to survey the archaeological excavation. The availability of LaserScanner 3D during investigation in the field raises many issues: firstly, how to manage the processing of the point-clouds (Paradata); secondly, how to distribute any kind of 3D document for further analysis and research (Metadata); finally how to share this knowledge using modern tools which exploit the Semantic Web (Ontologies).
The paper deals with these issues and gives an overview on the future perspective of the archeological documentation, integrating new software like Adobe Acrobat 3D with ontological representation as CIDOC-CRM for the structured and unstructured texts and GML (with some adaptations) for the management of spatial and geographical documents.

1. INTRODUCTION

For some time, in order to check the effectiveness and efficiency of 3D acquisition tools during excavation, CISA of the University L'Orientale of Naples has been experimenting with the LaserScanner to gather 3D archaeological documentation. The tests aim at outlining the scenarios which concern the applicability of a highly precise and analytical instrument to the stratigraphical method.

Experiences carried out so far regard the creation of three-dimensional models of archaeological excavations or urban and territorial investigations. These are dealt with tests which include problems of graphical acquisition and rendering: from stratigraphical explorations to surface reconstructions, from simple architectonic survey to complex monuments.

The research program is based on the definition of a methodological pipeline specifically for digital surveying (architectonic, structural, of the surface, of the findings) which can be integrated with data gathered by others specialists.

3D LaserScanner has been utilized for the rapid creation of highly accurate geometric surveys. In some cases, the survey of archaeological structures was possible only through the employment of laser scanning technologies: resorting to traditional techniques would have required long acquisition times and less precision.

Up to now, our experimentation has concentrated on the reconstruction of geometries and surfaces, allowing us to read and analyze the models and to create views, plans and sections. Besides the models are navigable at 360°, offering views that would not otherwise be possible to see. They also constitute valuable documentation of the archaeological object, which can be useful during the phase of geometric and material definition of the artifact, as well as in the preservation, monitoring and possible restoration phases.

The 3D LaserScanner, although it requires a long learning curve and currently has still elevated costs, is a technology which in the near future will be employed more extensively to document and study archaeological contexts.

The "technological" migration from an indirect digital survey that is completely unconnected from the interpretation of the remains raises two problems. The first concerns the competence and training of those who work in this field. The second concerns the need to guarantee homogeneity in the production and management of the geometrical information acquired, as far as the type of file produced (open and/or proprietary formats), the creation of the model (Paradata), and finally accessibility and reuse of the models (Metadata) are concerned.

Based on experience gathered in the field, this contribution intends to highlight certain trends in application, while pointing out the areas of "distress", which are particularly evident in the standardization of the procedures, therefore in the accessibility of the three-dimensional models.

Only an adequate circulation of these innovative digital sources can effectively favor the exploitation of informatics and its products. Already there are many experiences leading in this direction, aiming to define standards in order to guarantee the transparency of the 3D models, therefore their reproducibility.

After a brief introduction on the role of graphical documentation in the stratigraphical investigations and the synthetic description of a few case studies, this contribution will highlight the role that certain software assumes in scenarios of interoperability and reuse, software like Adobe Acrobat 3D, able to "enclose" the 3D models within normal PDF documents or modeling language for geographic systems, in particular GML and an adapted version of it called CityGML.

2. DOCUMENTING THE EXCAVATION

In the handbook on "Archaeology, Theories, Methods and Practice", Renfrew e Bahn (2006, p. 102) highlighted that a good excavation is characterized by the value of the archaeological documentation produced during the field work (drawings, plans, sections, photos, forms and reports) rather than in the extension of the investigations or in the marvel of the discoveries.

Judgment of an excavation is based on the meticulousness of the data produced in the course of the excavation, rather than on evaluating the reconstruction and/or interpretation of what has been destroyed in the course of the investigation.

Forms, photographs and drawings are the material objects of a strategy performed in the field by the archaeologist. They are the final products of a practical behavior carried out by observing an operative method and praxis.

For these characteristics the excavation can be assimilated to a standardized procedure. By agreeing with a specific workflow, this process anticipates various phases of intervention in which the connection between operations is rigorous.

Standardization not only allows to resort to a universally accepted code of regulations, representing a factor in scientific quality, but also to obtain a homogeneity in data acquisition. Forms, diaries, reports, photographs and drawings are the material witnesses to the excavation and, as such, inalienable elements within the "circumstantial evidence" paradigm.

The documentation procedure includes an articulated series of actions (material and immaterial) referring to converging activities (planning and evaluation; excavation, data treatment; communication). The nature and level of the data are generally dictated by specific and circumscribed requirements.

How does the change of methodology manifest itself, if archaeological documentation is comparable to a standardization procedure that includes the objectives and contents of the research, up to the excavation strategy and the rules used to formalize the data acquired?

Is it enough to modify one of the above mentioned elements (forms, photos and drawings) to declare that we have changed the "investigative" strategy?

The digital revolution has without a doubt radically altered the creation of those informative and testimonial which have characterized the practice of stratigraphical archaeological excavation for over half a century. We've gone from B/W photos and slides, to digital cameras reaching high levels of resolution; from simple hand drawings to sophisticated electronic equipment such as EDM, GPS, terrestrial photogrammetry and 3D LaserScanner. Finally, the excavation diaries, once written by hand, are nowadays compiled on PDAs and managed with increasingly more complex informatics systems.
Is this technological innovation sufficient to justify – as Kuhn (1978) put it – a change of paradigm?

A number of scholars, even though they understand the importance of computers, have never truly investigated the innovative role held by informatics in the field of archaeology. The most complete and exhaustive study on the function of informatics in field archaeology was undertaken by I. Hodder (2000). The scholar enumerates 4 out of 12 strategies dedicated to computers. Some strategies seem to be reproposed, according to the well-consolidated post-processualist opinion, an attitude which reduces informatics exclusively to technical tasks. Others are praised by Hodder, as computers are seen as innovative instruments.

If excavation is not a technical activity, but rather an on-going production of hypotheses and interpretations influencing the initial stratagem, sharing the documentation produced and exchanged digitally supplies, according to Hodder, continuous and rapid data transmission, providing a constant update of what is brought to light.

The uninterrupted flux of information can easily be guaranteed by the availability of data stored in a database. At the same time different types of information gathered on site (plans, drawings, object measurements, films and excavation diaries) can easily be encoded and made accessible to researchers. In Hodder's mind, data circulation represents the temporariness of the conclusions, which are always momentary. They become definite only at the conclusion of the procedure.

To guarantee constant and steady traffic of information, a certain degree of formalization is necessary. The user, in order to contextualize the records, may turn to other types of data. Stored and indexed excavation diaries are used, for example, as sources of the excavator's considerations and on his/her evaluations pertaining to the questions raised during on-site research.

According to Hodder, the dig must be filmed not only to document the main phases of the intervention on the terrain, but also of the fears, reflections and possible afterthoughts. Video recording completes the excavation diaries, supplying visible proof of the actual excavation.

In his hypotheses dedicated to informatics, Hodder highlights the technical, yet very important role that the computer has in guaranteeing data circulation, therefore the precise and rapid comparison of all the digital information available.

In our opinion it does not solely concern the managerial "improvements", according to the idea of efficiency that always accompanies the employment of computers in the archiving procedures. Massive use of informatics, even where data transmission networks are involved, not only ends up modifying the classification of finds, the interpretation of the stratigraphical units and the a posteriori reconstruction of the entire life of the area of interest, but the entire excavation progress.

Informatics isn't just a technical-operative means, but also a conceptual tool that constantly influences the researcher in the field, as it requires very standardized forms of normalization and data encoding. It is therefore a part of the excavation method.

As it isn't possible to separate the acquisition techniques from the end product of their application, we can maintain that the more widespread use of LaserScanning technologies providing graphical documentation of archaeological excavations, will determine profound alterations in excavation methodologies and procedures, just as in the geometric description of artifacts and in their subsequent analysis, classification and interpretation.

3. CASE-STUDIES

The equipment used to perform the work is produced by Zoller & Frohlich (http://www.zf-laser.com/e_index.html), a German company. The 3D scanner, Imager 5003, is a tool for short and medium range applications (minimum distance 40 cm. up to 53,5 m.), yet it guarantees elevated resolution (max. 36.000 x 15.000 pixel: horizontal by vertical) with a velocity of 500.000 pixel per second. The linear error specified by the company is less than 5 mm.

Three scanning profiles are available: Superhigh, High e Medium. At 10 meters distance the area sampling for High shoots is 6 mm, while for Superhigh it is double.

The maximum coverage of the area to scan is 360° x 310° (horizontal by vertical), needing about 6 minutes for a complete scan at High resolution. The Scanner, having an angle of 50° vertical not covered by the shoot (coinciding with the station's point of installation), generates a shadow cone, i.e. an absence of data, whose diameter varies according to the height of the tripod, generally positioned at 170 cm.

The Imager 5003 has an internal calibration system. The distance and angle calculation system is based upon the measurement of the phase. Data are acquired both in spatial coordinates x, y, z, and in reflectance values. The latter data, shown in grayscale with an interval between 0 and 255, corresponds to the material's "response" to the laser beam: thus it is possible to scan surfaces which are not illuminated.

For the post-processing phase we used the JRC 3D Reconstructor (www.reconstructor.it) an application dedicated to the process of point clouds. The JRC 3D Reconstructor is a software package used for processing laser data, 3D models and 2D images. It includes fast interactive pre-registration and automatic refinement of registration, multi-resolution meshing, flexible camera calibration and texture mapping, inspection and surveying tools (ortho-photo, cross section extraction, direct connection with AutoCAD drawing tools).

JRC 3D Reconstructor can also process Non-Structured Point clouds (point from airborne laser, from photogrammetric and topographical measurement - GPS total station), and shows some Projectors (orthographic, perspective, cylindrical, spherical) for real-time texturing. The independency from a specific sensor and the ability to re-sample the project data in various ways makes JRC 3D Reconstructor a 3D meta-processing software.

The surveys were carried out according to a consolidated pipeline which include planning the survey, the correct organization of the targets, and the subsequent merge of the scans, filtering and the elaboration of meshes (Scopigno, 2006).

We adopted a standardized and well consolidated procedure in acquiring and processing 3D data. The surveys were carried out analysing different types of archaeological monuments in order to test the laserscanning methodology in diverse conditions during excavations or fieldworks. The acquisition phase was a quite fast technical part: from one to four workdays; while the processing was a longer step in accordance with the archaeological needs, evaluations and interpretations. At the end of the first test campaign we found that the use of laserscanner had changed some attitudes and behaviours about work on field and the following data analysis. The availability of a rapid survey technology, able to gather spatial data in different environmental conditions, pushed archaeologists to acquiring more maps and drawings than in a conventional fieldwork with traditional equipments. A huge amount of processed data had therefore to be classified, archived and managed.

3.1 Pompei

Since 2004 the exploration of Insula 7, Regio IX has been underway, under the direction of Prof. F. Pesando, in order to understand both the settlement patterns of a sector of the city located outside the area commonly considered as the *Altstadt* and of the building history of single houses. So far investigation has concentrated on houses nr. 21, 23, 25 e 26, the north western corner of the block.

Research is based on documentation and on the survey of the existing housing structures. It is followed up by investigation in the field, needed to recover data regarding possible pre-existing or initial phases of occupation of the site.

The survey of the structures is particularly important because, thanks to accurate graphic documentation of the elevations, it is often possible to single out the different phases of construction of the buildings (vertical stratigraphy), which are no longer detectable at the surface or foundation level.

In September 2006, a 3D LaserScanner program to survey the façades and interiors of the housing in Insula 7 was carried out with the objective of defining a method to make the "extraction" and reading of vertical stratigraphy of the structures in elevation a semi-automatic procedure.

The first phase of the intervention focused on houses IX, 7, 19 and 21 along Vicolo di Tesmo and houses IX, 7, 22 – 23 – 25 and 26 along Via degli Augustali.

The campaign lasted a single day and affected a front of about 75 meters. The data, gathered into eleven different scenes, was acquired in "High" modality, thus guaranteeing sub-centimetre precision and with a resolution adjusted to the objective of the survey (fig. 1).

After the scans were recorded and assembled, a 3D model was created onto which several color images taken with a digital camera were rectified and projected. Some planes were subsequently generated to allow the creation of photo-orthoplan Then the phase involving vectorialization in CAD of the different building techniques was initiated.

The pertinent descriptive form, containing indications on the physical relationships between the various building techniques, will be associated to each type of masonry identified.

3.2 Wadi Gawasis

In January 2008, during a mission directed by Prof. R. Fattovich, an initial digital scan of two caverns was carried out; the laser scanner was used to guarantee the acquisition of detailed documentation of the caverns excavated in the fossil coral plateau and utilized probably as boat shelter from the beginning of the Second Millennium B.C. (Bard, Fattovich, 2007)

At the time of the intervention the caves presented problems related to statics, manifested by some collapses on the face and ruins near the entryway to cave 2, which had been buttressed with wooden reinforcements.

Figure 1: Pompei: three different shoots before the registration. In evidence the shadow cones.

The need to build an accurate geometric model that would constitute the basis for subsequent static monitoring of the structure, was added to the scientific need to graphically document the caverns' interiors. The survey lasted 4 days in the field and required about 4 weeks for post-elaboration in the lab.

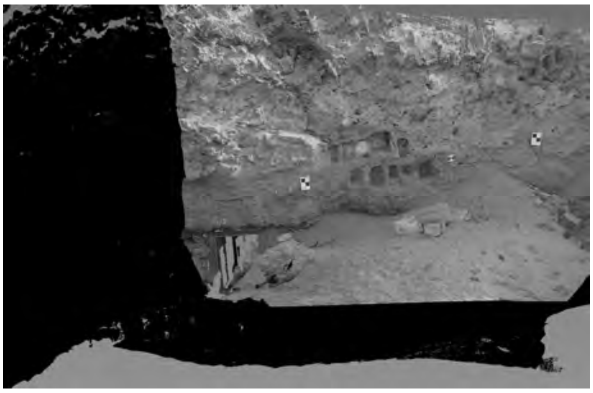


Figure 2: Wadi Gawasis: rectified Photo fitted to the point-cloud.

The survey was programmed and targets were affixed to the face of the cavern in order to allow subsequent reassembly of the scenes. Afterwards, more targets were placed on the interior walls of the caves with the same purpose. Particular care was dedicated to the placement of the intermediate targets in order to reassembly the three distinct surfaces (front, cavern 2 and cavern 3) into a single model. This allowed to rearrange and assemble all the point clouds into a single geometric reconstruction.

3D rendering of the area was later "coated" with textures taken from digital photos, opportunely calibrated and adjusted (fig. 2). The model was then used to extract important information, such as the section and plan of the cavern. At the same time the presence of deep lesions to the ceiling were detected, both in the section and in the plan.
The acquisition of new scans, expected in the next campaign expected in 2009, will allow to compare the reconstructed models and verify the progress of possible lesions, thus guaranteeing adequate monitoring of the monument's stability.

3.3 Temnos (Turkey)

At Temnos, in October 2007, during the last survey campaign performed in the framework of research activities lead by Prof. G. Ragone, a digital survey of a large portion of the terracing wall of the agora was undertaken. The tract of wall, in polygonal opera, was located on the western corner of the ancient Aeolian city.

Objective of this intervention, performed in a single day, consisted in the detailed documentation of the building techniques. Therefore a visible part of the wall of about 43 meters not covered with vegetation was chosen.

The LaserScanner was positioned about 4-5 meters from the wall, positioned on a slope slightly lower than the level of the monument, then 4 scans were acquired and elaborated

After the various scans were assembled, a mesh was created and on it were projected rectified images obtained with a digital camera. Subsequently a photo-orthoplan was created and exported to AutoCAD, where it was elaborated in order to extract the sections and to carry out the vectorial analysis of the polygonal masonry (fig. 3). The section clearly shows that the tract of wall was constructed against the earth, slightly leaning inwards to support the pressure from the terrace.

The area of our investigation is not easy to reach. It would have taken many days to perform the scans with traditional techniques. The laser scanner, on the other hand, allowed us to rapidly acquire the point clouds that were later opportunely elaborated in the lab.

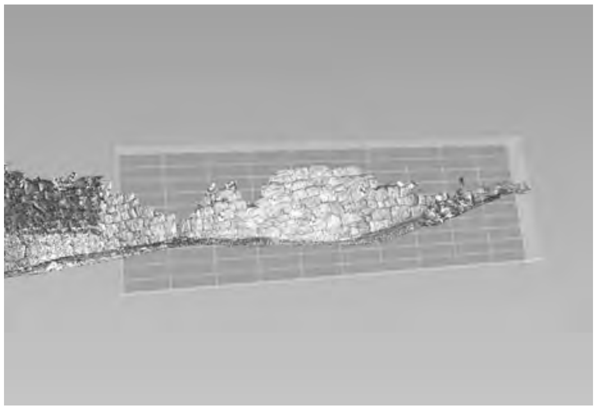


Figure 3: Temnos: the extraction of photo-orthoplan

4. USE AND REUSE OF THE 3D DATA

The availability of 3D data will determine an alteration of the investigative and archaeological excavation methodologies. Therefore the evaluation of, and attention placed upon, the techniques for acquisition and rendering of the three-dimensional models must be adapted to the complexity of the technology used in order to avoid ambiguity or – worse - inappropriate use.
The geometric basis of the digital model is an element to take into consideration. The third dimension is often not measured, but deduced from extrusion functions or reconstructions in perspective. Instead, a 3D reconstruction must be completely measured in all its components, the parts elaborated through modelling must be differentiated from those actually surveyed.

Another element must therefore be taken into consideration: the non selective property of 3D LaserScanner acquisition, along with the irrelevance of the factors of scale in graphic rendering.

All manuals on archaeological surveys, when examining indirect survey techniques, frame them mostly in a topographic kind of survey within a local or absolute reference system, although they're increasingly used to analyze and render details.

In traditional surveying (direct and indirect) the points to scan are chosen by the operator, whether he/she is an archaeologist, draftsman or architect. During acquisition he/she evaluates in a discriminating way the information to graphically record. The 3D LaserScanner memorizes all the data that the laser's optic beam can capture.

The digital master, obtained through a non-sampled acquisition of the archaeological evidence, can subsequently be elaborated: the model can be cleaned up, filtered and simplified. Only upon termination of the analysis will the spatial and physical qualities of the archaeological object be identified.

In a point cloud the concept of scale is irrelevant. The model is in itself scalable, as it is vectorial. The problem with scale presents itself only when rendering and printing the model, not when viewing it. The possibility to represent the models directly, without any kind of graphic mediation, constitutes a relevant factor in the procedures for using and reusing 3D data.

The more widely spread approach in the field of archaeology requires the description of archaeological objects almost exclusively in 2D, evidently in connection with the type of paper support on which such information is rendered. The same stratigraphical matrix, which records synchronic and diachronic events, is a simple bi-dimensional representation of the sequence of activities that the archaeologist has identified.

Recent attempts to intervene on the third dimension and on volumes (VOXEL) highlighted the geometric value of the isolated stratum in that over and under spatial relationship, which depicts the most interesting type of temporal relationships (before and after). (Cattani, Fiorini, Rondelli, 2004).

Is it possible to maintain the 3D geometrical dimension of the model without having to convert it into a two-dimensional view? Can we really make it more usable, without necessarily having to project it into a two-dimensional representation or without forcing possible users to install sophisticated and not very practical software for virtual navigation? Finally, is it possible to attribute relative semantic values to simple geometric superimpositions?

Starting from our on field experiences we tested some solutions (commercial or not) in order to store and manage all digital surveys. Our impression is that the Acrobat 3D software and the GML language could open new extraordinary scenarios for archiving and three-dimensional topological structuring of spatial data.

The functioning and the applicability of standardization and distribution of the 3D contents will be examined in the following paragraphs.

4.1. Acrobat 3d

Acrobat is a program by Adobe Systems used for creating and modifying PDF files (Portable Document Format). In the late Nineties it became the standard for the distribution of contents made up of text and images. The Adobe Reader software, freely distributed by Adobe, allows end-users to open, read and print the PDF files. PDF is a format based on a language for the description of pages, developed to represent documents autonomously from the hard- and software used to generate or manage these files.

The PDF does not include specific information, allowing it to be visualized and rendered the same way independently from the platform and/or device used to read it. This *portability* has permitted PDFs to become the *de facto* standard format, widely spread for sharing documents. In December 2007, PDF became an ISO 32000 standard.

In 2006 Adobe launched the Acrobat 3D program and in the Spring of 2007 the version v.8 was released, consenting to convert and compress CAD files into a single PDF 3D file readable with the free Reader 7.0 software. Acrobat 3D allows to create PDF documents from three-dimensional models obtained through various applications accurately maintaining the object's measurements. It is possible to visualize a model in all its aspects, having a wide choice of navigation and query functions at disposal. It won't be necessary to own CAD-like software.

Acrobat 3D supports bi-dimensional "static" models, organized in distinct informative levels, and navigable three-dimensional models: the user is able to rotate the model, measure distances, create sections, modify the illumination and rendering.

PDF is a web standard and for this characteristic it is used to distribute electronic texts. If the documents are not password-protected, they can be viewed within the browser through a special viewer; it is possible to download files in pdf format corresponding to specific requisites by adopting particular research techniques.

Documents incorporating 3D models with annotations and texts explaining the procedures used in the model's construction (Paradata) can be made available by exploiting the PDF's characteristics of portability.

Paradata (Scheuren, 2000) are a special category of observational data introduced in documentation process to more precisely determine specific features of the collecting/measuring/capturing of data. They are useful when we want to reuse manipulated data (not raw data), stored in the archives, or to have information about functioning and performance of processes generated. For this reasons Paradata may be described as a way to check and alter any process procedure evaluating costs/benefits in order to optimize the processing.

Figure 4 shows a simple example of annotation of paradata recording the software used to convert and merge the original GML model in a pdf file.

Linking Paradata with the 3D data should a positive approach to force researches to make transparent their processes according for instance the principles of LondonCharter (Beacham, Denard, Niccolucci, 2006). In the same time this special category of annotations could help other users to test and reuse the processes in order do not invent again the wheel.

Figure 4: A pdf file: 3D model with annotations.

By annotating the processes in PDF it will be possible to search Paradata online through the traditional search engines and so understand better the generated reconstructions.

Finally, by resorting to certain functions implemented directly in the software it will be possible to use metadata, also in DublinCore format, linked both to the document and to the single 3D objects incorporated in the PDF.

The creation of 3D PDF documents will improve the interoperability between different CAD, thus stimulating forms of collaboration and *networking*. Moreover, the introduction of texts and annotations explaining the construction procedure of the 3D model would greatly clarify how to access and reuse the three-dimensional object.

Today technology promotes major integration and forms of *peer production.* However it does not yet allow a broad management of the digital models, due to the "heaviness" of certain reconstructions. The development of lighter standard formats like X3D or COLLADA may guarantee in future the full interoperability of 3D documentation based on favoring a more widespread circulation of the digital masters and the relevant digital processing.

4.2. GML and CityGML

Acrobat 3D, by integrating CAD objects codified in 2 or 3 dimensions, favors the exchange of contents and information. However, to ensure that the data is correctly reused, a means acting as exchange is not sufficient; only the 3D data can be visualized in its geometric representation.

To incorporate semantics within the model, a language in the representation of the elements having a two- or three-dimensional spatial connotation is necessary. This objective cannot be reached by turning to other domain onthologies, such as CIDOC-CRM, which have nonetheless proven to be applicable to the archaeological documentation sector (D'Andrea 2006, D'Andrea, Marchese, Zoppi 2006). The expressive potential of GML (Geography Markup Language), but most of all of one of its application-oriented profiles denominated CityGML, must be analyzed. Through appropriate modifications CityGML can be used to examine the semantic relationships of ancient monuments.

GML is an XML grammar defined by the Open Geospatial Consortium (OGC) to express geographical objects. A standard since 2007, GML is a modelling language for geographic systems and an open format used to exchange geographic information through the Internet. GML defines *feature* an entity which is distinct from a *geometric object.*

A *feature* is an object of the application representing a physical entity (a building) that might not have a geometric aspect. A feature collection is a collection of features that can itself be regarded as a feature. As a consequence a feature collection has a feature type and thus may have distinct properties of its own, in addition to the features it contains.

A *geometric object* instead defines a location or region, rather than a physical entity. This distinction between the physical and geometric representation of a spatial object differentiates GML from GIS models, where *feature* and *geometric object* are considered equivalent.

The state of a *feature* is defined by a set of properties, where each property can be thought of as a {name, type, value} triple. The number of properties a feature may have, together with their names and types, are determined by its type definition.

Other relevant elements in the organization procedure of the *feature* are the *coverage* class, representing the discrete geographic coverage of the entity in a well defined spatial-temporal domain and the concept of *observation,* which models the observation and measurement action also defined according to precise spatial-temporal coordinates. The OCG, in the formulation of their guidelines, defines *observation* as a simple class, suggesting the development of specific GML application schema applicable to scientific, technical and engineering observation and measurements, such as the 3D LaserScanner.

CityGML constitutes an application profile for GML3, oriented at the multilevel representation of informative strata such as buildings. "CityGML" is a model oriented predominantly to the codification of three-dimensional urban objects.

In the example shown in figure 5 it is possible to understand how semantics is expressed through a simplified hierarchy based on node *parent-child* linked by the relationship *is part of.*

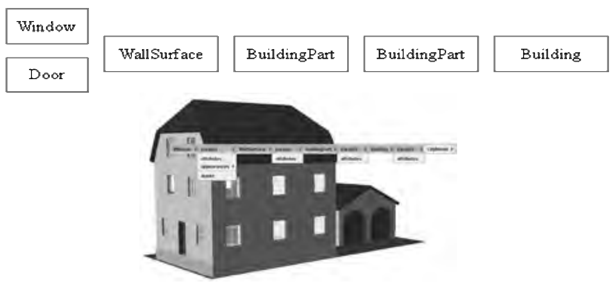


Figure 5: CityGML Schema.

Currently CityGML has been defined for the representation of urban contexts and geometries also to manage problems like traffic, pollution and for the simulation of catastrophic events (www.citygml.org).

Through appropriate modifications, the CityGML profile can be adapted to the archaeological domain for the semantic, not only geometric, representation of the constructed stratigraphical units. An ancient building is made up of a number of physical entities, such as interiors, roofs, doors, walls, floor area, foundations, furnishings, etc. Each physical entity has its own geometry and a series of attributes (dimension, chronology, building technique, etc.) which can easily be expressed, as in the example in figure 5.

The parts that are missing or that have been totally reconstructed can always be represented both geometrically with properties defining the degree of reliability and also according to fuzzy parameters (Hermon, Niccolucci, D'Andrea 2005).

The point clouds showing the ancient object's data are included in the GML grammar as *coverage*, i.e. the specific result of an *observation* which is the end product of a survey performed with a LaserScanner in a defined spatial-temporal ambit.

Representation of the stratigraphical units seems to be more complex, especially that of the negative type, as in the construction of the matrix they are the trace (or interface) of a "removal" event (natural or artificial), therefore devoid of geometry. The S.U. are always the result of an action which can be of contribution and removal; the former are generally Positive and Built Units, while the latter do not have a volume and therefore are Negative.

GML is, also in this case, the grammar used to conceptually express the S.U. Each stratum identified by digger is the trace with its own geometric properties (*geometry*) testifying an ancient event marked by an action of contribution or removal (*feature*). To highlight the chronological and functional sequence of the strata, the schema *topology_xsd* must be enriched and the physical relations which characterize the relationships between the S.U., currently circumscribed simply to the before-after and over-under sequence, should be anticipated.

There are many open-source and commercial GIS that elaborate documents written in GML (TatukGIS, OpenJump, QGIS, GVSIG and GRASS), while some programs are able to read files encoded in GML (GMLViewer). Managing documents in GMLCity is more complex, as they are viewable and correctly rendered only by specific programs such as Aristoteles and Landxplorer CityGML viewer.

Currently Reconstructor software can export point clouds in X3D and Collada format. We hope that a more coordinated action between producers and users will in the future contribute to the creation of interoperable formats like GML with the possibility to add and define properties and attributes of the recorded information.

The next step in this direction will be the integration of GML and its application profile to the CIDOC-CRM; GML could manage spatial components associated with the description of objects formalized according to CIDOC-CRM schema (Felicetti, Mara 2008). An alignment with the CIDOC has already been defined for X3D (Niccolucci, D'Andrea 2006). GML takes shape as a dominion able to express not only two-dimensional geography, but also the three-dimensional one. It is also able to record the topological relationships between *features*.

5. CONCLUSIONS

As laserscanning technologies are increasingly used to graphically document archaeological excavations, it will determine profound modifications in excavation methods and procedures, in the geometric description artifacts and in their subsequent analysis and classification.

If using new instruments to collect geometric data allows to rapidly gather highly accurate 3D information, it is very likely that the entire excavation and reconstruction procedure will gain significantly from it, in terms of quality of the data acquired and in terms of the completeness of the geometric-spatial information.

Notwithstanding this, a deep transformation of the methodology will occur not only thanks to more powerful and reliable digital technologies, but also when the product of excavation activities will be completely available to the entire scientific community. Conceptual and physical instruments already exist and are the infrastructure of the Semantic Web, dominion ontologies and standard formats. What is missing is the willingness of archaeologists, still entrenched behind circumscribed circles of knowledge, ever more restricted and auto-referential.

Until about 20 years ago, there was a limited number of discoveries and excavations and every archaeologist could, in a convenient temporal dimension, read and be informed of progress in his/her specific field of research. The extension of research, the extreme apportionment of knowledge, the appearance of a new phenomenon represented by local communities - in certain areas of particular historical interest in developing countries, excluded initially due to a sort of colonial attitude, even in the field of research – makes it difficult to keep the literature updated.

In August 2007 the Project Heathrow T5 on-line was launched, making available to users the alphanumerical and spatial data pertaining to excavations carried out at Heatrow and Stansted airports (www.framearch.co.uk). The system was developed thanks to a joint venture between Wessex Archaeology and Oxford Archaeology. The update system includes an improved version of their free GIS that allows you to explore the hard archaeological data collected during the phases of the investigations. Crucially, the raw data behind the GIS has been released in a variety of useful formats including CSV, XML, SHP and GML.

The diffusion of tools to simplify access to and reuse of the data will favor, thanks to the transmission of metadata and paradata, the creation of collaborative communities characterized by the peer production modality, the result of vast horizontal participant networking.

Achieving this objective is made possible not only because of a change in attitude, but also due to the existence of semantic and technological standards, and to a rigorous formalization of point clouds. The integration of standards, 3D representation languages and technologies constitutes a scenario from the near future.

Our tests – still partial - show how is possible integrate in an interoperable framework 3D data acquired by laserscanner and which directions should be undertaken in order to guarantee the access and the re-use of digital 3D data. This new approach for data archiving will improve the effectiveness of standard - at the moment only in the theory as regard 3D data - and could support different European digital libraries projects and international archives for spatial and geographical data.

6. REFERENCES

Bard, K., A., Fattovich, R., (Eds.), 2007. *Harbor of the Pharaohs to the Land of Punt. Archaeological Investigations at Mersa/Wadi Gawasis Egypt, 2001-2005*. Università degli Studi di Napoli L'Orientale, Napoli.

Beacham, R.; H. Denard and Niccolucci, F., 2006. An Introduction to the London Charter. In: Ioannides, M., Arnold, D., Niccolucci, F., Mania, K. (Eds.), *The e-volution of Information Communication Technology in Cultural Heritage: Where Hi-Tech Touches the Past: Risks and Challenges for the 21st Century*. Project papers from the joint event of VAST/CIPA/EG/EuroMed 2006, Budapest, Archaeolingua, p. 263-270.

Cattani, M., Fiorini, A., Rondelli, B., 2004. Computer applications for a reconstruction of archaeological stratigraphy as a predictive model in urban and territorial contexts. In: Ausserer, K.F., Börner W., Goriany, M., Karlhuber-Vöckl, L. (Eds.). *Enter the Past: the E-way into the Four dimensions of Cultural Heritage. CAA 03, Computer Applications and Quantitative Methods in Archaeology*, Proceedings of the 31st Conference, Vienna, April 2003, pp. 299-303.

D'Andrea, A., 2006. A preliminary Ontology-based model applied to the description/interpretation of the archaeological excavation. In: von Hahn, W., Vertan, C. (Eds.), Proceedings of the First International Workshop on "*Ontology Based Modelling in The Humanities*", 7-8 April 2006, University of Hamburg, Bericht 264, pp. 38-46.

D'Andrea, A., Marchese G., Zoppi, T. 2006. Ontological Modelling for Archaeological Data. In: Ioannides,M., Arnold, D., Niccolucci, F., Mania, K. (Eds.), Proceedings of the 7th International Symposium on *Virtual Reality, Archaeology and Cultural Heritage (VAST2006)*, Nicosia, Cyprus October 30 – November 4, Switzerland, pp. 211-218.

Felicetti A., Mara H., 2008. Semantic Maps and Digital Islands: Semantic Web technologies for the future of Cultural Heritage Digital Libraries. In: Kollias, S., Cousins, J. (Eds.). *First Workshop on Semantic Interoperability in the European Digital Library,* 2nd June 2008, Tenerife, Spain, pp. 51-62.

Hermon S., Niccolucci F., D'Andrea A. 2005, Some Evaluations on the Potential Impact of Virtual Reality on the Archeological Scientifich Research. In: Thawaites, H. (Ed.). Proceedings of the Eleventh International Conference *on Virtual Systems and Multimedia*, VSMMS2005, October 3-7, Ghent, Belgium, Budapest, pp. 105-114.

Hodder, I., (Ed.), 2000, *Towards Reflexive Method in Archaeology: The Example at Çatalhöyük*, McDonald Institute for Archaeological Research. British Institute of Archaeology at Ankara, Monograph No.28.

Kuhn, T. S., 1978. *La struttura delle rivoluzioni scientifiche*. Einaudi, Torino.

Niccolucci, F., D'Andrea, A., 2006. An Ontology for 3D Cultural Object. In: Ioannides, M., Arnold, A., Niccolucci, F., Mania K. (Eds.), Proceedings of the 7th International Symposium *on Virtual Reality, Archaeology and Cultural Heritage (VAST2006)*, Nicosia, Cyprus October 30 – November 4, 2006, Switzerland 2006, pp. 203-210.

Renfrew, C., Bahn, P., 2006. *Archeologia. Teoria, metodi, pratica.* Zanichelli, Bologna.

Scheuren F.J., 2000. Macro and Micro Paradata for Survey Assessment. Paper presented in a satellite meeting to the *UN/ECE Work Session on Statistical Metadata*. Washington D.C., U.S. Bureau of Labor Statistics.

Scopigno, R., 2006. Gestione efficiente dei dati prodotti dai sistemi di scansione tridimensionale. In: Campana, S., Francovich, R., (Eds.). *Laserscanner e GPS. Paesaggi archeologici e tecnologie digitali*. All'Insegna del Giglio, Firenze, pp. 41-68.

7. ACKNOWLEDGEMENTS

The author wishes to thank Professors F. Pesando, R. Fattovich (University of Naples L'Orientale) and G. Ragone (Univeristy of Rome 3) for having graciously allowed the publication of the case studies. The graphic elaboration of the point clouds was performed by Giancarlo Iannone. Many thanks to Giancarlo Iannone, Laura Saffiotti and Matteo. Lorenzini for the discussions on 3D and on spatial tagging languages for 3D elements. To Stephanie Williams a special thanks for the revision of the translation.

LOOK OUT! CAN YOU SEE SOMETHING COMING? SEMANTIC BROWSING – WHAT MIGHT IT LOOK LIKE?

K. May [a], C. Binding [b], D. Tudhope [b]

[a] English Heritage, Strategy Department, Fort Cumberland, Portsmouth PO4 9DE UK.
- Keith.May@english-heritage.org.uk
[b] Faculty of Advanced Technology, Glamorgan University, Trefforest, Wales.
cbinding@glam.ac.uk - dstudhope@glam.ac.uk

KEYWORDS: Ontology, CIDOC CRM, CRM-EH, RDF, SKOS, Semantic Web, STAR

ABSTRACT:

The Semantic Technologies for Archaeological Resources (STAR) project has been exploring the development of new tools and technologies for "semantic web" based research. The project builds upon the ontological modelling approach taken by English Heritage staff, in recent years, to modelling their information and data using the CIDOC-CRM standard (ISO 21127:2006). The ontological modelling has enabled the EH archaeological teams to more explicitly identify where 'information gaps' exist within their existing information records and flow lines, which can then be bridged using an ontological information model. The data involved can be derived from legacy datasets, current databases, and hopefully will enable incorporation of a mapping to data yet to be recorded in a newly implemented archaeological recording system. The aim is to provide a model for how new systems and technologies can be developed that enable greater interoperability and better integration of data in the rather disparate archipelago of archaeological project information.

This paper will also discuss the emerging use of the Simple Knowledge Organization System (SKOS) data model as a W3 standard for sharing and linking knowledge organization systems via the Semantic Web. The STAR SKOS web services currently provide term look up across the thesauri held in the system (including the EH Monuments Thesaurus and the former MDA Objects Thesauri), along with browsing and semantic concept expansion within a chosen thesaurus. Users may browse a concept space to explore and become familiar with specialist terminology or as part of a broader application. In combination with a search system, the services allow queries to be expanded (automatically or interactively) by synonyms or by expansion over the SKOS semantic relationships. Expansion is based on a measure of 'semantic closeness'.

The paper will also introduce a new prototype CRM Browser web service developed by the STAR project and explore ideas for how this prototype browser might be developed further in the future to enable linked searching across and between free text reports and structured data in databases, using emerging forms of semantic query languages and interfaces and what such "Semantic Browsing" might look like for end users.

1. INTRODUCTION

1.1 Overview and structure of this paper

This paper will first recount some of the background to why the ontological modelling work was carried out by English Heritage archaeological staff and how this was carried forward by the Semantic Technologies for Archaeological Resources (STAR) project. It will then give a short summary of the conceptual modelling carried out and discuss in more detail the use of extensions of the CIDOC CRM to give greater definition of the main archaeological concepts involved. The paper will present an overview of the methods used by the current STAR project for mapping from the CRM-EH model to specific archaeological data sets, including the use of the emerging W3C standard SKOS (Simple Knowledge Organisation System) for implementing controlled terminologies in ontological modelling systems. The paper will then consider some of the current issues involved in formulating, gathering, representing and reflecting user requirements for a semantic browsing system, as yet to be produced. Finally there will be some consideration of the future directions for possible semantic web browsing and searching, followed by overall conclusions.

1.2 Background and aims of the Ontological Modelling

The English Heritage Revelation project (May, S., 2004) identified early in its assessment stage that EH was not lacking in archaeological information systems. Rather the picture was of an archipelago of self-contained and isolated islands of information that had been designed over the last twenty-five years or more. Most of these EH archaeological systems were designed to fulfil individual project requirements, but without the overall planning and structure to enable the shared use and interoperability of the data being collected or created.

The data flow diagrams and entity relationship models of the existing archaeological systems helped to give a clearer picture of the baseline state of affairs. But the resulting systems documentation was still a series of rather fragmented data models for each system without a clear method for how best to integrate the data held within each of them. It was decided to attempt to model both the existing information holdings but also include further new information requirements that would be wanted in a newly designed system. This was intended to express more explicitly where the gaps were, both in and between the existing data models, and most significantly showing where those gaps might be filled or "bridged" by modelling new relationships between those bits and pieces of information. At that point attention was drawn to Semantic Web developments and in particular the CIDOC CRM (Crofts et al. 2003) and solutions that might be provided by an ontological approach to data modelling.

More of the background to the ontological modelling of the English Heritage (EH) archaeological information domain has been presented at the Computer Applications in Archaeology conference in Prato 2004 (Cripps & May 2004, forthcoming) and further publications and outputs are available from the CIDOC CRM website including an online version of the model and accompanying documentation (Cripps et al 2004, May, 2006). Since the beginning of 2008, and in this article, the English Heritage Conceptual Reference Model (CRM) has been referred to as the "CRM-EH", to distinguish it from the CIDOC CRM ontology from which it derives, and which it is still directly related to.

1.3 Background and aims of the STAR project

STAR is a 3 year Arts and Humanities Research Council (AHRC) funded project, in collaboration with English Heritage and the Royal School of Library and Information Science Denmark, applying semantic and knowledge-based technologies to the digital information of the archaeological domain. The project aims to develop new methods for linking archived and 'live' digital databases; associated vocabularies; and, where relevant, related grey literature, exploiting the potential of a high level core ontology (CRM-EH) and natural language processing techniques.

Increasingly within archaeology, the Web is used for dissemination of reports and the associated datasets that result from fieldwork or scientific analysis of material from historic environment investigations. This contributes to the growing amount of information on the 'deep web', which a recent Bright Planet study estimated to be 500 times larger than the 'surface web'. (Bergman, 2001). However Google and other web search engines are ill equipped to retrieve information from the richly structured databases that are key resources for humanities scholars. A higher and higher proportion of recent archaeological results and reports are appearing as grey literature, increasingly online, before or instead of traditional publication. Typically these are not indexed or made available for searching other than as ordinary web documents. It is difficult using conventional search engines to link these to datasets or indeed to search them using terminology other than that employed by the authors. Cultural heritage and memory institutions generally are seeking to expose databases and repositories of digitised items - previously confined to the realm of specialists - to a wider academic and general audience. The mapping from lay (or related subject area) terminology to technical vocabularies in a particular domain is a critical problem. There is a need for new tools to help formulate and refine searches and navigate through the information space of concepts used to describe a collection. Different people use different words for the same concept or may employ slightly different concepts and this 'vocabulary problem' is a barrier to widening scholarly access.

The historic environment sector has a rich tradition of employing Knowledge Organisation Systems (KOS – such as thesauri). However, such vocabulary tools are often not fully integrated into searching and indexing systems and online practice has tended to mimic traditional print environments. The full potential of these knowledge resources in online environments has not been tapped. The paper will explain how the emerging W3C standard - Simple Knowledge Organisation System (SKOS) - can provide the necessary semantic cross-referencing for term look up across the thesauri held in the system, and enable browsing and semantic concept expansion within a chosen thesaurus. This will allow a search to be augmented by SKOS-based vocabulary and semantic resources (assuming the services are used in conjunction with a search system). Users may also simply browse a concept space to explore and become familiar with specialist terminology or as part of a broader application. In the next section, this paper will consider how the use of an ontology such as the CIDOC CRM, especially when further enhanced by domain specific extensions (CRM-EH), can provide semantic interoperability between previously isolated datasets built using different database platforms and designed with differing data structures.

2. CONCEPTUAL MODELLING USING CIDOC CRM

The CIDOC CRM standard (Crofts et al., 2008) does not require (nor particularly recommend) any particular methodology for using it. After consultation, the approach that was adopted by the CFA was derived from general ontology building methods (Denny, 2002). and can be summarized in five main stages (Cripps, 2004):

- Acquire domain knowledge
- Organize the ontological model
- Flesh out the ontological model
- Check the work
- Commit the ontological model

In the first instance this resulted in a model that related archaeological conceptual classes directly to the CIDOC CRM entities. However on further consideration it was found that the scope notes within the CIDOC CRM that were the definitions of the semantic meanings behind the concepts, only represented the archaeological concepts in the CRM-EH model at quite a high level of conceptualisation. It was for this reason that further specifically archaeological extensions, with more specific archaeological scope notes, were made (Cripps et al 2004).

3. ARCHAEOLOGICAL EXTENSIONS OF CIDOC CRM

In a recent paper D'Andrea succinctly summarizes one of the key problems with any attempts to integrate the digital information recorded by archaeologists, namely the nature of archaeologists themselves: "Undertaking a standardization process involving archaeologists and archaeological data may perhaps be considered as a symptom of naivety. Few scientific communities are more individualistic than this, the result being an extreme fragmentation of systems and data models." (D'Andrea 2008). Nevertheless, the work undertaken by EH staff to develop better forms of integration for their own archaeological project work in recent years, does in some way attest to an equally strong professional ethic of expecting, and feeling professionally obliged, to share the information outputs from their historic environment recording work, both with other archaeologists, but perhaps even more so with the wider public. It was largely because of the recognition of the fundamental problems represented by the diversity of different archaeological recording systems developed by the many professional archaeological organisations in England over the last twenty five years (May, S., 2004), that English Heritage staff decide to look for an approach that would allow archaeologists to "map" their existing systems to some common and over-arching information framework, rather than trying to re-invent yet another "recording system to end all recording systems".

As a result of investigating the possibilities for a common information framework the CIDOC CRM was considered as a possible solution for bringing such a plethora of archaeological information datasets together, and building the semantic links by making the semantic relationships explicit – between the various islands in the archaeological information archipelago. However, although the CIDOC CRM was derived from the wider cultural heritage domain it was noted during the early stages of modelling the archaeological activities of EH staff, that not all the entities in the 'vanilla' CIDOC CRM were quite explicit enough to cover some of the more complex relationship. In consultation with the CIDOC CRM Special Interest Group it was decided to *extend* the entities that were represented in the EH ontological model (CRM-EH) and in order to distinguish them and their semantic meanings an additional numbering system was produced. Most importantly additional 'definitions' amounting to scope notes pertaining to the more specific archaeological entities represented by the CRM-EH entities have been written and these are now contained in the RDF 'Description' field as part of the RDF structure that forms part of the CRM-EH model. Thus a CRM-EH entity ContextFind (EHE0009) is also a CIDOC CRM entity Physical Object (E19). The CRM-EH extensions and scope notes are principally derived from EH archaeological experience, practice, and in particular the relevant information already documented in the EH archaeological recording manual, upon which the majority of any digital recording system is necessarily based. By effectively "mapping" the CRM-EH to the fields in the EH recording system the CRM-EH has to some degree been 'future-proofed' to make sure that it relates as closely as possible to any computer system that is implemented in the future to contain the data that the paper based recording manual is also meant to record.

This is why the CRM-EH is referred to as an *extension* of the CIDOC CRM – although the provision for such extensions is an integral part of the CIDOC CRM's implementation as stated on the opening page of the CIDOC CRM: "By its very structure and formalism, the CRM is extensible and users are encouraged to create extensions for the needs of more specialized communities and applications". (Crofts, N., et al 2008 p i). This paper will show how the use of the CRM-EH extensions has further aided the process (if not the simplicity) of mapping between different archaeological datasets, although it is noted that it is not an absolute requirement to make extensions to the CIDOC CRM in order to map data and indeed in many cases within the STAR implementation a choice has been made to simply map directly to 'vanilla' CIDOC CRM without any requirement to use a CRM-EH extension. At present it remains to be seen exactly what further issues of interoperability may emerge as more and more archaeological data sets are mapped, and more and more complex searches and queries are attempted. As D'Andrea notes: "Considering some of these mapping procedures, it may be noticed that there are alternative ways of representing the same conceptual archiving practice. While the English Heritage mapping chose to base on the creation of new sub-classes of "IsA" type specializing the original CIDOC-CRM and making it richer, the Italian ICCD mapping preferred to maintain a full compatibility with CIDOC-CRM, fitting the starting source with the destination ontology only through the semantic equivalence between corresponding classes". It should be clarified here that the CRM-EH is fully compatible with the CIDOC CRM – indeed every CRM-EH entity also bears the related CIDOC CRM entity and "E" number. What the extensions have allowed is greater 'specialisation' of specific issues pertaining to how EH archaeologists (and probably many others) actually record the archaeology that they investigate. More details of the RDF extensions and associated scope notes can be found by downloading the RDF files from:
http://hypermedia.research.glam.ac.uk/media/files/documents/2008-04-01/CIDOC_v4.2_extensions_eh_.rdf

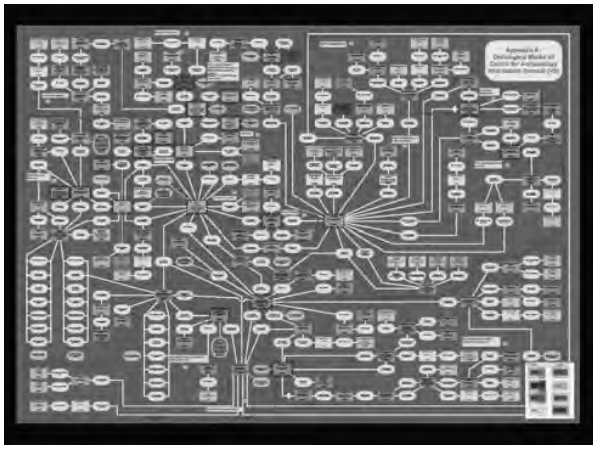


Figure 1: Ontological Model of the English Heritage Archaeological Information Domain
See also online version at
http://cidoc.ics.forth.gr/docs/AppendixA_DiagramV9.pdf

4. STAR INITIAL METHODOLOGY FOR MAPPING FROM THE CRM-EH MODEL TO DATA SETS

This section will set out some of the main issues addressed in establishing a methodology for mapping from the CIDOC CRM ontology and CRM-EH extensions to actual specific data fields in various selected test-bed datasets.

4.1 Dataset selection

To test the possible implementation, and prove the interoperability of the modelling, a number of data sets were needed. The test-bed data were selected by various criteria: data from EH legacy systems such as Delilah (itself over 25 years old); current or more recent databases being used by teams in the EH archaeological centre at Fort Cumberland, including some specialist archaeological science data from the Environmental Archaeology Branch at EH (mostly a variety of different project databases using various versions of MS Access or specialist datasets in MS Excel); external data in a database designed by a different archaeological organisation not using EH software or systems; and some data that reflected recent developments in online publication of integrated data and reports. The LEAP Silchester data, structured using IADB (http://www.iadb.org.uk/), and published in Internet Archaeology as an example of "Integrated Publication" (Miles, 2004, Richards 2004) satisfied both the latter criteria. In addition some attempt was made to choose datasets that broadly covered a range of archaeological periods, but ones that would allow some meaningful archaeological cross-project searches.

The initial datasets chosen were Raunds Roman Archaeological Database (RRAD) along with Raunds Prehistoric data, Raunds environmental sampling data and the Silchester LEAP data. As well as being from quite differing database origins these data sets were also from different stages in the historic environment project management process which archaeological projects tend to follow (English Heritage 2007): Raunds prehistoric data was the excavation data as archived after work on the site was completed; Raunds environmental data derived from the specialist environmental assessment work carried out by staff of the former Ancient Monuments Lab at English Heritage (Campbell forthcoming); RRAD is approaching the Analysis stage following on from the recommendations in the Assessment stage work; Silchester LEAP data was integrated with a 'fully' published and peer reviewed journal article in Internet Archaeology (Clarke, A et al. 2007). These characteristics of the initially selected four datasets are summarized in Table 1 below.

	Database Type	**Main Archaeology Periods**	**MoRPHE Project stages**
Raunds Prehistoric	Delilah - CSV	Neolithic & Bronze Age	Execution - Excavation
Raunds Environmental	MS Excel - DBF	Neolithic to late Roman	Execution - Assessment
Raunds Roman (RRAD)	MS Access - MDB	Roman & Iron Age	Execution – Assessment to Analysis
Silchester LEAP data	MySQL - MYD	Roman & Late Iron Age	Execution - Publication

Table 1: Summary of initial test-bed data sets for STAR prototype CRM browser

4.2 Data mapping and generation of RDF triples

In order to map the datasets to the CRM-EH (and thereby to the CIDOC CRM) the approach taken was to identify a "core" of key archaeological concepts from the larger CRM-EH data model and then relate these "core" entities to the key fields identified in the RRAD, RPRE and IADB databases. From this starting point further data fields could be mapped to the CRM(s), as and when they needed to be included in the resulting merged test-bed RDF dataset. This intellectual mapping required 'domain' archaeological knowledge of the data and the CIDOC CRM and CRM-EH ontologies. Initial mappings were performed by May and communicated via spreadsheets to the team in Glamorgan. The process of 'domain mapping' is time consuming, and requires considerable focus on the complex semantic and conceptual issues being addressed. It is not therefore something that can be easily slotted in to a few minutes here and there within a general work schedule. Fortunately it seems that once an initial mapping is produced for the archaeological domain, the process of mapping a further system dataset is considerably aided by being able to recognize similarities (or exact matches) in work-patterns or conceptual activities and then simply using the same relevant parts of the CRMEH model. From the initial "core" mappings, it did generally become easier for some subsequent mappings to be performed by others in the project team (i.e. non archaeological domain experts) using the initial spread sheet mappings as a guide, with domain expert validation by May as and when required.

In addition some of the exchange of the mappings between entities and core data fields and, particularly, keeping the evolving CRM-EH modelling up to date, was aided by the use of Protégé ontology editing software (http://protege.stanford.edu/) or Altova SemanticWorks (http://www.altova.com/products/semanticworks/semantic_web_rdf_owl_editor.html) although in many cases the complexity of the modelling diagram was beyond the graphical visualization capabilities of these primarily text based ontology editing programmes. One more recent development of the modelling is an attempt to add more dimensions and "granularity" to the modelling diagram with the use of multiple colours and colour shading, but this also has drawbacks for displaying greater complexity on any computer screen, not least it's inaccessibility for colour-blind readers. The further process of actually extracting the data from the datasets in accordance with the "core" mappings, was also time consuming and would not be human scaleable over a large number of datasets. Therefore an automated data mapping and extraction utility - using SQL queries with query parameters saved in XML format for subsequent reuse - was developed by Binding to assist the processing of the end data, and the resulting output is an RDF format file. The automated mapping utility consists of a form allowing the user to build up a SQL query incorporating selectable consistent URIs representing specific RDF entity and property types (including CRM, CRM-EH, SKOS, Dublin Core and others). The extracted data was imported into a MySQL RDF triple store database on the Glamorgan server, using the SemWeb RDF library.

4.2 ID format adopted

The RDF entities in the RDF triple store, require unique identifiers. Some of the data being extracted was an amalgamation of records from separate tables – e.g. *EHE0009.ContextFind* actually contained records from RRAD.Object & RRAD.Ceramics tables. It was therefore necessary to devise a unique ID for all RDF entities beyond just using the record ID from an individual table. The format adopted to deal with all these issues was a simple dot delimited notation as follows:

[URI prefix]entity.database.table.column.ID
e.g. "EHE0008.rrad.context.contextno.100999"

This format (although verbose) allowed the use of existing DB record ID values without introducing ambiguities. In RRAD database, Ceramics and Objects were both instances of *EHE0009.ContextFind*. This therefore involved the combination of data from two tables:

- *EHE0009.rrad.object.objectno.105432 [an EHE0009.ContextFind record from the RRAD object table]*
- *EHE0009.rrad.ceramics.ceramicsno.105432 [an EHE0009.ContextFind record from the RRAD Ceramics table, with the same ID value]*

The format also allowed the same base record ID to be used for both *EHE0009.ContextFind* and

E*HE1004.ContextFindDepositionEvent* (these records actually originated from the same table and had a 1:1 relationship), using a different entity prefix to disambiguate the records:

• *EHE0009.rrad.object.objectno.105432 [The ContextFind record ID]*
• *EHE1004.rrad.object.objectno.105432 [The ContextFindDepositionEvent recordID]*

Finally an arbitrary URI prefix (http://tempuri/) was added to all ID values. According to need, this can be replaced with a more persistent prefix.

4.3 Date/Time and coordinate formats adopted

Although there is nothing dictated in the CIDOC CRM ISO or CRM-EH about date/time representation formats, it was important to maintain a consistent date format throughout the merged data. For the purposes of the data extraction to keep all data consistent we used a "big endian" format (i.e. from most to least significant) compatible with both W3C standards and ISO8601 ("Data elements and interchange formats – Information interchange – Representation of dates and times"). The format is as follows:
CCYY-MM-DDThh:mm:ss e.g. "2007-05-03T16:19:23"
This format does not introduce any restrictions on how dates & times are eventually displayed or used within applications; it merely provides a common string representation mechanism for interoperability of data.

Spatial co-ordinates with various formats are used in a number of different fields in each of the tes-bed datasets. RRAD coordinates were 6 digit numeric values in separate "Easting" and "Northing" columns. RPRE coordinates were slash separated string values, sometimes with an extra 4 digit value appended (i.e. either *nnnnnn/nnnnnn/nnnn* or *nnnnnn/nnnnnn*). IADB coordinates were numeric values in separate "Easting" and "Northing" columns (and appeared to be relative to a site local reference datum). CRM/CRM-EH requires a single string to represent a spatial co-ordinate value. The consistent format chosen for output was 6 digit space delimited Easting and Northing values, with an optional Height value (Above Ordnance Datum). These values were all assumed to be in metres:
nnnnnnE nnnnnnN [nn.nnnAOD] e.g. "105858E 237435N 125.282AOD"

4.4 Modelling notes/annotations

The EH recording manuals and the current datasets contain several kinds of note fields. For the purposes of disambiguating all the different types of notes that show up in the RDF triples, a core set of EH archaeological note types have been identified. These are:
• Comments (the most general category of 'catch-all' notes)
• Method of excavation
• Interpretation (likely to be further refined to specific cases)
• Siting description (reasons relating to location of a sample)
• Site treatment (relating to samples)

While it might potentially be restrictive to model notes as strings (notes have other implicit attributes such as language, author/source etc.), this is the current position within the CRM (*E1.CRM Entity _ P3.has_note _ E62.String)*. However, taking the RDFS encoding of CIDOC CRM recommendation, we intend to create sub properties of *P3.has_note* e.g. *EHPxx1.has_InitialInterpretation*, *EHPxx2.has_RevisedInterpretation*, as part of future work. The CIDOC CRM has a modelling construct in the form of "properties of properties". For example, property *P3.has_note* has a further property *P3.1.has_type* – intended to model the distinction between different types of note. Unfortunately, this construct does not translate well to RDF. As evidence of this, property *P3.1.has_type* is not actually part of the current RDFS encoding of CRM on the CIDOC website (in the comment header there is a suggestion to create specific sub properties of *P3.has_note* instead). The more recent OWL encoding of CRM also avoids including the construct.

4.5 Modelling of Events

Many of the events defined in the CRM-EH modelling and used to interconnect objects and places in the CRM-EH model were largely only *implicit* within the earlier relational database structures. The ability to model these events more *explicitly* for further data recording and analysis was one of the key reasons the CIDOC CRM was chosen for the event based ontological modelling, allowing a new approach to systems modelling and design. As a result therefore, in certain cases during the mapping from the data to the CRM model (and its further translation from relational data structures to an RDF graph structure) it was necessary to create this *explicit event* information by the formation of intermediate '*virtual entities*', but with no current data to actually fill such entities. The aim is that for any newly implemented digital recording system used by EH, these explicit events will be recorded as part of those new recording systems. Being a higher level conceptual model the CRM has little intrinsic provision for the representation of actual data instance values. The approach adopted for the STAR data extraction process was to create rdf:value relationships as an additional property to model instance data for entities wherever appropriate.

4.6 Initial mapping of data fields to extended CRM

When imported into the SemWeb MySQL triple store database the combined data files produced the following results:

Database	**Entities**	**Literals**	**Statements**
RRAD (inc. STAN)	919,017	126,691	2,383,216
RPRE	114,105	20,482	317,085
IADB	85,694	21,592	209,582
LEAP	30,066	7,954	78,122
Totals:	**1,148,882**	**176,719**	**2,988,005**

Table 2. Statistics for extracted data

The number of statements (triples) contained in the resultant RDF files is **2,988,005.** Some triples (e.g. *rdf:type* statements) were duplicated due to entities occurring within multiple files, but any duplication was removed during the aggregation process. A number of separate RDF files were combined in the aggregation process including the CRM itself, the CRM-EH extension, alternative language labels for the CRM, and various EH domain thesauri. This scale of triple store has proved to provide perfectly adequate browsing capabilities for the project team even when using a relatively remote 3G wireless connection to the data on the Glamorgan server from a fieldwork site and during a mainline train journey. The data files produced were each validated against the W3C RDF validation service. Whilst this did not prove the validity of the data relationships or even conformance to CRM-EH, it did at least give confidence in the validity of the basic RDF syntax.

4.7 Prototype Search / Browse Application

An initial prototype client application has been produced (see Fig 2), capable of cross searching and exploring the amalgamated data in the RDF triples store. The application utilizes a bespoke CRM based web service for all server interaction (the underlying SemWeb library does also support SPARQL querying). Boolean full-text search operators facilitate a measure of query refinement and result ranking. Retrieved query results are displayed as a series of entry points to the structured data; it is then possible to browse to other interrelated data items, by following chains of relationships within the CRM-EH, and also working upwards from data items to concepts as desired.

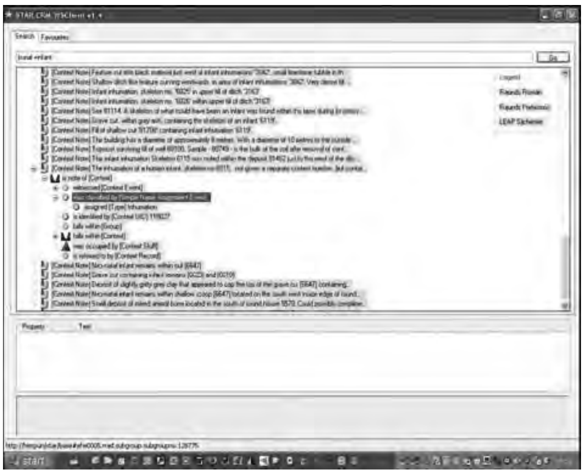


Figure 2: Prototype CRM Browser Web Service API

The results to date have successfully demonstrated the potential for searching across at least four very separate datasets with quite different origins and a variety of different data structures and content. At present the types of searching that can be carried out are inevitably limited by the constraints of the actual datasets selected for test purposes – although even with these constraints some very interesting archaeological queries have already been formulated that could not easily be explored without the use of the STAR browser. Further testing is planned to introduce additional datasets for query generation and expansion testing purposes, but also to test the practicalities of how much data the triple store can function with. How this CRM semantic browser could be developed in order to best enable archaeologists to search the various historic environment datasets will be discussed further in section 6.

5. THE USE OF SKOS BY STAR

The STAR project has adopted the emerging W3C standard SKOS Core as the representation format for historic environment domain thesauri, such as the English Heritage National Monuments Thesauri (http://thesaurus.english-heritage.org.uk/), and related Knowledge Organization Systems (KOS). "In general, thesauri conforming to the SI/NISO/ISO standards should map in a fairly straight forward manner to SKOS. However, there may need to be judgments on how to deal with non-standard features (Tudhope 2008). The STAR investigations to date have encountered some potential problems associated with the use of Guide Terms or facet indicators in some thesauri. Other issues may arise from the need to create URIs for concept identifiers as part of the conversion and the potential for validation.

6. ATTEMPTS TO GATHER USER REQUIREMENTS FOR A SEMANTIC BROWSER.

With the CRM Browser operational the next task has been to try and gather ideas for how archaeologists (as the CRM-EH domain experts) and other more general users may wish to use the capabilities to search across a range of different data sets. One immediate problem in judging this is that users have very little, if any, experience of what "semantic browsing" might look or 'feel' like. In the first instance a number of workshops are being held to gather 'user requirements' in a fairly orthodox manner. These have already led to a discussion of possible interfaces for searching in systematic ways between finds and contexts and navigating around groups of contexts to understand better all the different types of data that are associated semantically with a single unit of archaeological Stratigraphy and then working up these queries across associated archaeological groups to search across an associated group of archaeological features (in different datasets) and further still to answer the sort of complexities posed in the 'scenario' "Excavated examples of 'Roman' corn drying ovens where archaeobotanical analysis has taken place and where rye (seacale cereale L) is found in deposits from the 'Iron Age' in Europe" or "What is the first record for Centaurea Nigra (Black Knapweed) in England"? Another area of interest for archaeologists will be the potential for geospatial browsing using GIS, (something not particularly within the scope of the current STAR project although some exploration in this area will be attempted).

The current CRM Browser is just a prototype of further tools that the technologies will enable to be developed. At the present time, more use of the cross-terminology searching enabled by SKOS will also need to be developed and one use of the SKOS web service for archaeologists is likely to be a web service that allows archaeologists to better specify what particular archaeological terminology list they are referring to when they deposit their digital archives or write their archaeological reports. One further style of user interface that has been postulated, is the development of more interactive methods for users to select from a suite of 'tools' for semantic searching, perhaps allowing a spatial, temporal and text based series of search mechanisms to be available in the "semantic query toolkit".

6.1 Future directions for linking reports to data

One major area that STAR is still investigating is the ability to use the CIDOC CRM and CRM-EH ontologies to aid the semantic searching of reports linked to data and visa versa. In particular there seems some potential to use Natural Language Processing (NLP) techniques such as those employed in the GATE architecture (http://gate.ac.uk/) to annotate reports with commonly identified archaeological concepts. This could thereby allow both authors and researchers to develop new approaches to formulating conceptual 'threads' through the information written about an archaeological investigation, and enable researchers and other users to 'drill down' into underlying data once they have identified a good 'semantic match' using a suitable semantic query interface, be that a basic text-based query ("more like *this* please?" whatever ***this*** represents); or GIS based (spatial - show me *where* these things come from?); or some sort of time-line (temporal – *when* did these sorts of things first start occurring?); a photo or even part of a photo (image based "give me more that look like this bit of the image"?)… or in the future most probably some combination of all these.

7. CONCLUSIONS

The results to date from the combined work of the ontological modelling and the outputs from STAR have demonstrated that a degree of semantic interoperability has already been attained. While this is encouraging and is a positive driver for further research, it remains to be adequately assessed whether a full implementation of the CRM-EH modelling is achievable in a cost-effective manner. This is in part because a full implementation of the CRM-EH model is only going to be really achievable once a new system for digital recording has been introduced at EH, in the next year, that will enable staff to begin to collect data for all the new entities in the model (and include data in those fields that accordingly map to the 'virtual entities') and it will then take some further time for enough examples of projects with this new type of related data to be recorded and analysed, to test the effectiveness of the new interoperability and how well it integrates the new data with the legacy data. Nevertheless the degree to which at least partial interoperability between a range of different datasets has been achieved, across at least three different types of database software and archaeological recording system structures, shows that the methodology is sound and only by the further testing of more and more archaeological datasets – mapped to the CRM-EH, and thereby to the CIDOC CRM – will reveal increasing tests and measures of the scalability of the current system and hopefully will enable the most appropriate ways to implement scaleable solutions for both hardware and software infrastructures. Work on the early stages of using NLP software for annotating key relationships within reports has also looked promising and if the same level of 'mapping' across different organisational report structures can begin to be attempted then it may be possible to incorporate many of the 'core' elements of the CRM-EH ontology into some form of developmental report and data search and browser mechanism.

7.1 Acknowledgements

The STAR project is funded by the UK Arts and Humanities Research Council (AHRC). The views expressed here are principally our own, but they are based on the wider work of the STAR project team and domain experts at English Heritage.

7.2 References

Bergman, M., 2001. WHITE PAPER: The Deep Web: Surfacing Hidden Value. *Journal of Electronic Publishing*, vol. **7**, no. 1, August, 2001 (accessed 16 June 2008)

Clarke, Amanda., Fulford, Mike G, Rains, Mike and Tootell, Klare., 2007. Silchester Roman Town Insula IX: Development of a roman property C. AD 40-50 - C. AD 250. *Internet Archaeology* 21. http://intarch.ac.uk/journal/issue21/silchester_index.html *(accessed 28 May. 2008)*

Cripps P., Greenhalgh A., Fellows D., May K., Robinson D., 2004. Ontological Modelling of the work of the Centre for Archaeology, CIDOC CRM Technical Paper. http://cidoc.ics.forth.gr/technical_papers.html

Cripps P., May K., (forthcoming). To OO or not to OO? Revelations from ontological modelling of an archaeological information system. *Proceedings of Computer Applications in Archaeology* Prato 2004.

Crofts, N. Doerr, M. Gill, T. Stiff, M. and Stead, S. (Eds.) Nov 2008. *Definition of the CIDOC Conceptual Reference Model and Crossreference manual. Version 4.2.4.* Official release of the CIDOC CRM. (Accessed 9 June 2008)| http://cidoc.ics.forth.gr/official_release_cidoc.html

D'Andrea. A., Niccolucci, N., 2008. "Mapping, Embedding and Extending: Pathways to Semantic Interoperability. The Case of Numismatic Collections". First Workshop on Semantic Interoperability in the European Digital Library **(**SIEDL 2008).

Denny, M., 2002. *Ontology Building: A Survey of Editing Tools.* http://www.xml.com/pub/a/2002/11/06/ontologies.html

Doerr, M., 2003. *The CIDOC CRM – an Ontological Approach to Semantic Interoperability of Metadata.* AI Magazine, Volume 24, Number 3 (2003).

English Heritage., 2007 *Management of Research Projects in the Historic Environment (MoRPHE).* (Accessed 16 June 2008) http://www.englishheritage.org.uk/MoRPHE

May, K., 2006 "Integrating Cultural and Scientific Heritage: Archaeological Ontological Modelling for the Field and the Lab". CIDOC CRM SIG Workshop, Heraklion 2006 http://cidoc.ics.forth.gr/workshops/heraklion_october_2006/may.pdf

May, S., et al 2004 *Revelation: Phase 1 Assessment.* English Heritage Research Report 78/2004

Miles, D., 2004 Digital dissemination and archiving. *Internet Archaeology* **15**. http://intarch.ac.uk/journal/issue15/preface (accessed 16 June 2008)

Richards, J., 2004 Online Archives. *Internet Archaeology* **15**. http://intarch.ac.uk/journal/issue15/richards_index.html (accessed 16 June 2008)

SKOS: Simple Knowledge Organization Systems, http://www.w3.org/2004/02/skos

STAR Project: Semantic Technologies for Archaeological Resources, http://hypermedia.research.glam.ac.uk/kos/star

Tudhope, D., Binding, C., May, K., 2008 Semantic interoperability issues from a case study in archaeology. First Workshop on Semantic Interoperability in the European Digital Library **(**SIEDL 2008).

May, K., 2006 “Integrating Cultural and Scientific Heritage: Archaeological Ontological Modelling for the Field and the Lab”. CIDOC CRM SIG Workshop, Heraklion 2006 http://cidoc.ics.forth.gr/workshops/heraklion_october_2006/may.pdf

May, S., et al 2004 *Revelation: Phase 1 Assessment.* English Heritage Research Report 78/2004

Miles, D., 2004 Digital dissemination and archiving. *Internet Archaeology* **15**. http://intarch.ac.uk/journal/issue15/preface (accessed 16 June 2008)

Richards, J., 2004 Online Archives. *Internet Archaeology* **15**. http://intarch.ac.uk/journal/issue15/richards_index.html (accessed 16 June 2008)

SKOS: Simple Knowledge Organization Systems, http://www.w3.org/2004/02/skos

STAR Project: Semantic Technologies for Archaeological Resources, http://hypermedia.research.glam.ac.uk/kos/star

Tudhope, D., Binding, C., May, K., 2008 Semantic interoperability issues from a case study in archaeology. First Workshop on Semantic Interoperability in the European Digital Library (SIEDL 2008).

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

NEEDLE IN THE HAYSTACK: FINDING, INDEXING AND SEMANTICALLY QUERYING IMAGE COLLECTIONS OF ANCIENT COINS

M. Sifniotis[a], A. Felicetti [a]

[a] VAST Lab, PIN, University of Florence (maria.sifniotis,achille.felicetti)@pin.unifi.it

KEY WORDS: Semantic search, Ontologies, Image Searching, Illegal Coinage Trade

ABSTRACT:

The COINS (Combating On-line Illegal Numismatic Sales) EU project aims to assist the fight against illegal trade of ancient coins. Its purpose is to retrieve catalogue and index coin images from the web and compare them by using image processing algorithms to known lists of stolen coins. COINS consists of three interconnected areas: a search engine that crawls the web and retrieves candidate images, an image recognition application for matching candidate coins with stolen ones, and a management tool for archiving, indexing modifying and querying a coin image archive. This paper provides research results achieved in two of those tools; the search engine and the management tool

1. INTRODUCTION

1.1 Background

Online illegal trade and selling of antiquities and art works is unquestionably a very important issue for authorities to handle. The COINS project aims at providing a substantial contribution to the fight against illegal trade of ancient coins which appears to be a major part of the illegal antiques market. It is an EU funded project which includes technological, heritage and law enforcement partners.

COINS is not a single piece of software; it provides a set of tools that can assist the user with their search for ancient coins. Three interconnected areas are available: a search engine that crawls the web and retrieves candidate images, an image recognition application for matching candidate coins with stolen ones, and a management tool for archiving, indexing modifying and querying a coin image archive. These tools are able to work independently as well as collaboratively.

This paper provides research results achieved in two of those tools; the search engine and the management tool. It is structured in five sections. This section provides an introduction and literature review on electronic illegal trade, web image crawling, aspects of semantic search and ontologies as well as the major contributions of this work. Section 2 describes the methodology used for the design of the two tools. Section 3 examines preliminary results while Section 4 outlines a conclusion and future work aims.

1.2 Electronic illegal trade

With the increasing use of the internet, online trade in stolen goods has reached a large scale. Research by CheckMend, an online stolen property checking service indicates that roughly £5 billion worth of stolen goods are on sale at any time in the UK. Its checking service identifies two items per minute as stolen – a total of 10% of all the second hand items checked with a projected value of £100 million (CheckMend, 2008). High profile cases of online illegal trade have surfaced in the news and retailers in the US are putting pressure on online marketplaces to safeguard the origin of the items being sold (Wutkowski, 2007). In the area of antiquities, illegal trade is also very prominent and the market is tentatively estimated as being worth around $2 billion a year; reliable data is hard to find (Conklin, 1994, p. 4). Once more, high profile unmaskings of illicit trade have taken place with notable example the £25 million case of 10,000 'unprovenanced' antiquities (Todeschini, 2007) recovered by the Italian carabinieri.

Illegal electronic sale of antiquities has not been actively documented in terms of volume but its existence is undoubted. In 2006 PAS (Portable Antiquities Scheme), the UK government funded scheme that records archaeological objects found by the public formed an alliance in 2006 with the online auction house, eBay, in order to curb unreported antiquities trade. In June 2008, the Swiss authorities made a deal with eBay to crack down on the illicit sale of cultural property over the internet. This will take the form of a three-month trial starting in July 2008.

Facilities such as CheckMend are very useful in cases where a serial number of an item exists or at least some sort of uniquely identifying property. On the contrary, identifying antiquities is considerably more difficult and the majority of successful discoveries until now have occurred because people chanced to look more closely at the items being sold. More appropriate to the case is trace.com, the global database of lost, stolen and seized items, which includes antiquities among others and is regularly updated by the public and law enforcement agencies. However, trace.com does not actively search for content on the Internet but rather relies on the vigilance and feedback of its users.

In the case of ancient coins, research by Elkins (2007) suggests that about 260,000 and 280,000 coins are sold each year on the eBay-U.S. website, not counting bulk lots.

1.3 Web image searching

Section 1.2 has illustrated that online trading of illegal antiquities is an active and ongoing problem. The COINS project is following an image-finding oriented approach. This means that we actively search the web for coin images.

Searching for images is still in its infancy stages. Quality of an engine's image results directly depends on the quality of the textual information associated with the images (e.g. filename, nearby text, 'alt' tags within etc). Advanced search abilities can display images according to colour properties (grayscale/colour) or size. In a few cases (Google, Exalead) the engines apply face recognition technology in order to identify likely images containing faces. Content-Based Image Retrieval (CBIR) search engines are few and mostly in research stages Examples include WebSeek and IBM's QBIC. These engines consider the characteristic of the image itself-its shape and colours.

Without the ability to examine image content, searches must solely rely on textual metadata. Queries on a CBIR system, except from the image's characteristics, can be by example (an image is provided by the user as a search criteria), or by semantic retrieval (finding images of a specific subject).

Recent research by Google (Baluja & Jing, 2008) attempts to blend computer vision techniques with user feedback and preferences. Images are returned by their visual similarities by using local features descriptors. Results are encouraging, showing a large decrease in the number of out-of-context images returned. This also follows work done for the National Centre for Missing and Exploited Children (NCMEC). In this scenario, Google assists NCMEC in tracking child exploitation and search for patterns in images of abuse on the web (Baluja, 2008)

1.4 Semantic search and ontologies

Semantic web

On a global scale, the Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is an initiative of the W3C consortium (W3C, 2008) implemented as a set of layered specifications. The core components are the Resource Description Framework (RDF) Core Model, the RDF Schema language and the Web Ontology Language (OWL). Interaction between these core components is provided by SPARQL (SPARQL, 2008), a standardised query language. This approach allows:

- Data to be surfaced in the form of real data, so that a program doesn't have to strip the formatting and pictures and ads off a Web page and guess where the data on it is.
- People to write (or generate) files which explain—to a machine—the relationship between different sets of data. This allows machines to follow links and hence automatically integrate data from many different sources.

As a result, semantic queries are queries that go beyond the simple inclusion of a keyword and the use of a search tool for navigation. A user of a semantic search can describe what he/she is looking for according to an object's properties. In other words, while in a classic 'navigational' search the result would be a single document, in a semantic search, the result would be information on a particular object, represented by a collation of data that may be retrieved across varying sources/documents (Guha, McCool, & Miller, 2003).

Ontologies

Previous research done by the authors in the framework of the EPOCH (EPOCH, 2008) and AMA (AMA, 2008) projects had as its final goal the use of open technologies and international standards to encode every aspect of the cultural heritage digitalization process. Results of this research (EPOCH, 2007) suggested that integration, preservation, sharing and distribution of cultural heritage information can be benefited by the conceptual richness provided by ontologies. An ontology-driven approach also couples well with semantic search.

The reference ontology used was CIDOC-CRM, a conceptual model specifically designed for the description of the cultural heritage world (Croft, Doerr, Gill, & Stead, 2008). The semantic richness provided by the classes and properties of this ontology allowed us to establish a conceptual layer able to describe all the entities that any user working in the cultural heritage field can deal with, and to integrate them in a logical and meaningful way. The tools developed with this in mind are AMA Mapping, a tool able to integrate information stored in legacy archives by mapping it to CIDOC-CRM, and the MAD (Mapping Archaeological Data) Semantic Container, a tool created to store, query and manage CIDOC-CRM encoded information and to expose it on the semantic web.

This approach was one of the first implemented in the area of cultural heritage. These tools are in the process of being adopted by the CIDOC-CRM committee in order to promote the use of semantic ontology among cultural heritage users (Felicetti A., 2006). The modular design of AMA/MAD allows them to be easily adapted in various cultural heritage scenarios. As a result, they have already been applied in Digital Libraries management and integration (Felicetti & Mara, 2008), and semantic enabled geographic systems (Felicetti & Lorenzini, 2007). Building on this work, we have been extending the toolset for the COINS project.

1.5 Contributions of this work

This paper makes three contributions:

1. We introduce an image-searching plugin to an open source search engine, able to retrieve both HTML pages and candidate coin images.
2. We introduce a metadata tool that is able to enhance the content retrieved by the search engine according to users' requirements.
3. We introduce an ontology-based semantic database capable of storing and querying information on ancient coins.

2. METHODOLOGY

The COINS management tool is the 'glue' of interaction between the different solutions offered by COINS such as the web spider, semantic engine, image recognition tools, and the thesaurus application. It works by providing a friendly set of interfaces for the user to interact. The final version of the tool will be released as a web application and it will be written using W3C standards.

We expect the system to be used by two different user types: numismatics experts and law enforcement. The numismatics experts would use it to define the coins reference collection, a set of selected coins to be used as a digital reference - an ideal coin collection for coherent sets (e.g. Etruscan coins) where optimal samples can be referred by humans or by machines for coin classification. This allows numismatists to build a complete coin knowledge base and thesaurus, by sharing and integrating their coin archives and their knowledge by exposing them to the semantic web. The law enforcement is provided with a powerful tool to discover and identify stolen coins illegally sold on the internet base on the numismatists' information.

Figure 1 illustrates the design of our interconnected tools. In section A, the Coins spider has the role of retrieving and saving appropriate content on its database. The content includes the HTML file, images, and any other metadata created on-the-fly in index time. The mapping tool (section B) provides the interaction between the spider and the management tool firstly by translating the retrieved metadata in a semantically meaningful way and secondly by giving the ability to add richer information to the content. Lastly, the Management tool (section C) among its other capabilities allows for navigational and semantic search of our data.

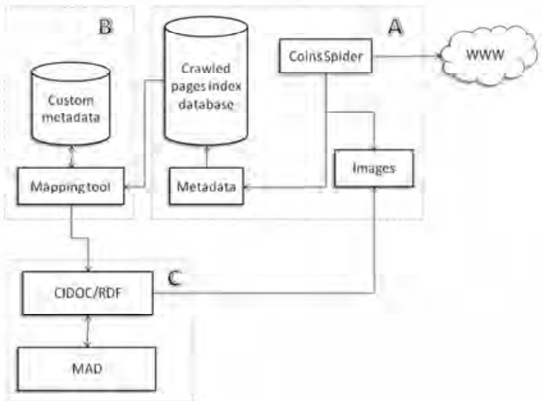


Figure 1. Diagram of tools and data flow

2.1 Search engine issues

When designing the search engine, we came across a number of factors which influenced our methodology:

1. *Whole web crawling:* In order to compare candidate coin images with a list of known stolen ones, good candidate sites must first be found. It is a tremendous task to crawl the whole web and requires a huge amount of hardware and man-hours.
2. *Open source approach:* One of the COINS project initiatives is its open source approach. This extends also to the search engine.
3. *Image indexing and caching:* The search tool should be able to index and save images as well as text.
4. *Resource constraints:* effort should be placed to limit the image candidates even when indexing the web: the less 'false candidates' the better.
5. *Metadata retention:* any information related to a specific image should be kept as potential information to be used by the management tool
6. *Multilingual properties:* Not all pages are written in English. The search engine should take under consideration the existence of multilingual websites.
7. *Index updates:* How often should an index be updated? How feasible is it to remove or updated outdated content?

Whole web crawling: It was clear from the initiation of the development that deep (extensive) web crawling in the count of millions of pages would be infeasible. Initial research was focused in identifying good candidate web sites containing ancient coinage. After consultation with the heritage experts of the project we are testing the web crawling with online auction websites. As a result, the crawler begins with a set of starting URLs and fetches relevant pages back. Of course, more starting URLs can be added in the process.

Open source approach: In an effort to be as open about our code as possible, we have opted to use an open source search engine implementation and build our image searching capabilities on top of it. Four open search engines were evaluated for the purposes of the project; their summaries can be found in Table 2.

	Web Glimpse	ht://Dig	SWISH-E	Nutch
Key Feature	Suffix arrays	Simplicity	Metadata	Ranking Excerpts
License	Nonprofit use	GPL	GPL/LGPL	Apache
Active	No	No	Yes	Yes
Crawling	Local FS	Intranet	Intranet	All
Caching	No	No	No	Yes
Clustering	No	No	No	Yes
Link Rank	No	No	No	Yes

Table 2. Comparison of open source search engines

Evaluation results indicated Nutch (2008) as the best candidate for our case. It is a complete open source search engine, implemented in Java and can be fully customised by programming plugins for it. Nutch is based on Lucene, another open source project for indexing and searching documents. Research shows that Nutch achieves high performance (Cafarella & Cutting, 2004) often as good as commercial search engines (Benedict, 2004).

Nutch is roughly composed of three parts: fetcher, parser and indexer (Figure 3). The fetcher is responsible for retrieving web pages from the web as well as updating any renewed content. The parser navigates through the fetched HTML (or other types) pages and indentifies tags and content available for indexing. Lastly, the indexer assembles indexable parts into the Nutch database. Nutch can also scale up or down according to the user's needs.

Its support for clustering uses Hadoop which implements Google's map/reduce computing paradigm and a Distributed File System (DFS).

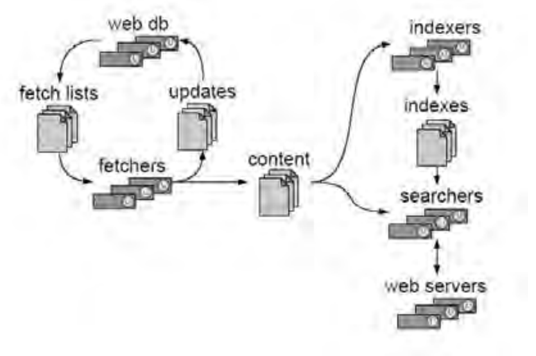


Figure 3. Nutch crawling/indexing procedure

Its use as a search engine is quite wide spread. Notable examples include Krugle (search for code, archives and technically-interesting content), Creative Commons (searches for Creative Commons licensed material) and the Internet Archive (searches National Library of Australia, the Bibliothèque Nationale de France, the Library of Congress, and the U.S. National Archives and Records Administration).

Image indexing and caching: By default, Nutch is not an image indexer. It works as any text-based search engine. Work such as (Zhang, 2006) introduce an image plugin able to read JPG file formats. However it has been tested on a local filesystem scenario and not on the web.

The solution was to design an image indexing and caching extension to the parser component. The plugin inherits all the properties of the text-based Nutch parser and extends it for the inclusion of images. As a result, this also aids us in the case of web page updates.

Resource constraints: An obvious solution to handling resource constraints is to limit the amount of pages the crawler is allowed to index. However, this approach may erroneously exclude positives. Additionally, it would be important to exclude automatic non-candidates such as banners and logos. The crawler is not aware of what it indexes; it does not know the content of a picture. Taking these under consideration, we have used an approach where candidate URLs are fully indexed, but images with size less than a certain threshold (90x90 pixels - customisable) are automatically ignored and not cached. This manages to exclude not only banners and logos but also thumbnail pictures that are too small for image comparison.

Metadata retention: By default, Nutch stores the whole textual content of a web page to its database. However, as the text is being retrieved, we can specify additional fields in the database that can dramatically improve the relationship with the management tool. Profiles of auction sites can be specified during the indexing and thus automatically include coin properties such as dimensions/mint type/etc.

Multilingual properties: To extensively test for multilingual capabilities we are including non-English URLs of auction sites.

Index updates: Updates to the index can be done as often or as infrequently wished. Different candidate URLs can be re-indexed in varying time intervals. For example, Auction site A can be queried for updates every 5 days while site B every one month. Initial experimentation has shown that a 25-day limit is good enough, since the majority of auction sites keep their ended listings data for at least one month.

2.2 Interoperability: the mapping tool

During the crawling operations it is important to find not only the image itself, but also the information concerning it, including the source from where the image is coming and the other metadata concerning it.

Once this data is harvested, it needs to be available for interaction and querying by the management tool. Additionally, the flexibility to add extra parameters to a retrieved image should exist. As a result, a mapping tool was designed to serve as an interaction platform between the spider's index and the underlying management tool.

2.3 The semantic encoder

The semantic enrichment of crawled information is done after the crawling operations. Every time new information is found in the spider's database, the mapping tool checks the spider's index, and encodes metadata information coming from images' web pages in a CIDOC-CRM compliant format on the fly. During this phase a XSL Stylesheet is dynamically applied by the mapping tool to the harvested metadata for an immediate encoding in RDF format and the creation of a set of RDF triplets,s describing entities and relationships of information concerning crawled images.

RDF triplets are then stored in a semantic container and ready to be queried and integrated with other semantic information coming from the COINS project frameworks (i.e. the CIDOC-CRM RDF encoded information coming from converted coin archives).

3. IMPLEMENTATION

3.1 Search engine

3.1.1 Case study-eBay

As mentioned in the introduction, a large proportion of illegal online trade occurs on the eBay site. As a result, the COINS spider uses eBay for testing purposes.

For this first case study we are using eBay US and specifically the ancient coins category. A study was made of four eBay sites, namely the French, UK, Italy and US ones for:

- The number of results in the ancient coins category
- The number of coin results in other categories.

Table 4 shows the number of results retrieved in May 2008 for all items in the Ancient Coins category, both for a local and a worldwide search. The US site gives the maximum number of results, which justifies its choice for our primary case study

Site name	Local search	Worldwide search
Italy	1686	6495
US	3383	7680
UK	935	6358
France	1388	4227

Table 4. Comparison of localised eBay sites

.Following this, we examined whether coins were found in other categories than the Ancient Coins one and whether these categories should be included in the crawl. Table 5 and 6 show the results of a simple keyword case; in Table 5 queries were done in the local language. Results are included from Ancient coins and the other categories. It was found that a very low number of coins appeared in other categories than Ancient Coins, mostly in the Antiquities one.

Keyword: Denarius	**Category: Coins**	**Category: Other**	**%**
UK	144	3	2.1
US	308	4	1.3
Italy	427	3	0.7
France	17	0	0

Table 5: Comparing results across categories (A)

Keyword: Roman coin	**Category: coins**	**Category: other**	**%**
UK	413	70	16.9
US	1026	74	7.2
Italy	114	5	4.4
France	91	3	3.3

Table 6. Comparing results across categories (B)

It would seem that eBay users try to be as accurate as possible when assigning categories for their listings. These observations also include items that may belong in both categories and would be thus retrieved from a crawl in just the Ancient Coins one. As a result the cost of crawling other categories for this case study was considered unnecessary.

Lastly, the COINS spider adheres to the good practices of web robots and fully respects robots.txt permissions before crawling a web page.

3.1.2 Overview of the COINS spider

As mentioned in 2.1 the COINS spider is based on Nutch. The crawler starts with the URL of eBay US Ancient Coins category and proceeds to fetch the links included. Aspects such as depth (how many links should be followed) and maximum number of links to be followed per page can be customised.

Experimentations with the eBay site provided a best depth of 3. This depth follows the top page, its links and the links on the fetched pages. More depth than this included a large number of erroneous entries while less did not give accurate results. The main reason behind this is the way eBay results are structured. While the specific category listing has its own domain (coins.listings.ebay.com) all referred items on eBay begin with cgi.ebay.com.

This was resolved by creating a profile set for the specific auction site. To illustrate, when the spider fetches a page from cgi.ebay.com it examines whether it belongs to the ancient coins category or not. If it does, it includes a customised tag in the database index that the specific page is a coin candidate. These profile sets can be extended to other candidate sites.

3.1.3 Image plugin

The image plugin is an extension of the default Nutch HTML parser. Apart from the usual HTML tags it also checks for `<img>` ones. After the fetching cycle finishes, the indexing commences, where data from the web sites is actually indexed in the Nutch local database. The plugin works as follows:

1. Find all image tags in the HTML file
2. Check if the image URL has been examined before by looking at the hash table – if it has, ignore it.
3. For each new image URL:
    a. Create a unique MD5 hash key composed by the URL of the image
    b. Add the hash key to a unique hash table
    c. Check if the image is valid and fits our dimension requirements
    d. If it does, save the file with its MD5 hash key as its file name.
    e. Include the MD5 hash file name as a tag inside our database index
    f. Include a property as a coin candidate
4. Repeat for next image

This approach solves a number of problems. Firstly, with the hash table it avoids re-examining the same file over and over thus saving considerable bandwidth strain. Secondly, it automatically gets rid of small thumbnail images as well as logos and banners. Thirdly it provides a meaningful way to retrieve all image candidates at once when a user is searching the contents of the database. The plugin works with all popular image formats such as JPEG, GIF, PNG, TIFF and BMP.

3.1.4 The search tool

Once the crawling is complete, the user is able to interact with the results by accessing the actual search. The search runs as a Java servlet on an Apache Tomcat server. The user can search by keywords and results displayed include also the images related to the specific page. Another option is to view only the images that have been retrieved from the crawling and classified as possible coin candidates.

3.2 Management tool: MAD: a Semantic Web Database

3.2.1 Overview

A powerful container is needed to manage the complexity arising from the data encoded using ontologies and the RDF syntax. It must deal with semantic data, just as relational databases deal with data stored in tables and records and with their relationships. A Semantic Database is needed for this, a big container where semantic information is stored and maintained to provide users with the necessary tools for their semantic search. For this purpose we created MAD (Managing Archaeological Data), a framework originally designed to manage structured and unstructured archaeological excavation datasets encoded using XML syntax, including free text documents marked up in XML.

The latest release of MAD comes with a multipurpose semantic engine able to store and manage ontology encoded information, i.e. data structured in CIDOC-CRM compliant format, a semantic query set of interfaces based on SPARQL and RQL query languages and a Firefox plug-in implementing a semantic browser for RDF graphs.

MAD can be used to store, browse and query semantic data in many powerful ways, but also to transform and supply semantic information on demand. The whole framework has been developed as part of an EPOCH activity for creation of information management systems for the Semantic Web and is entirely based on Open Source software, XML standards and W3C technology.

3.2.2 The MAD Semantic Web Database

MAD is built around a powerful Java native XML Database providing technology to store and index XML documents in a file-system-like structure of folders and sub-folders (collections). This container can be used to store information coming from mapped and digital metadata and content, along with annotations and tag sets created by users, mapping templates, schema, ontologies and everything can be expressed using the RDF language. Users can browse and query this integrated archive to get semantic information on Cultural Heritage objects described herein, or references to remote digital objects stored elsewhere (i.e. URIs and URLs linking specific resources). In this sense MAD acts as a Semantic Digital Library.

Semantic Queries in MAD

In order to query RDF data we are using the two most important languages available at present: SPARQL and RQL; two semantic query languages designed for retrieving information from RDF graphs. These languages provide clear, easy query syntax and the ability to obtain information from encoded documents without knowing its explicit syntactical structure (i.e. elements and properties names).

RQL is used due to its ability to combine schema and data querying using advanced pattern-matching facilities. SPARQL and RQL combined have been used for the creation of a group of semantic query interfaces able to parse the RDF documents in different ways. The ability of RQL to query ontology structures make the retrieval of classes, subclasses and properties from the models very simple allowing the building of structure-based semantic queries.

Classes and properties can be clicked to visualize subclasses and sub-properties and to define the elements that will participate to the query definition. While clicking on the different elements, an RQL query is constructed and then submitted, to be evaluated by the Semantic engine (Fig. 2). A human readable version of the query is also shown to make it understandable. An RQL query serialized on the CIDOC-CRM structure, for instance, may appear like this:

```
select $class0
from  {instance : $class0} @p {value : $class1}
where $class0 in subClassOf(
kyme:E28.Conceptual_Object )
and @p = kyme:P70B.is_documented_in
and $class1 in subClassOf( kyme:E31.Document )
and value like "*1022*"
```

It is the corresponding machine readable version of the clearer spoken language request: “I am looking for a *E28.Conceptual_Object* which *P70B.is_documented_in* the *E31.Document* containing '1022' “. Figure 7 illustrates the Semantic Query Interface created to allow users to build queries based on the combination of CIDOC-CRM classes and properties with free text.

Management Tool query capabilities

Numismatics professionals have a number of ways to query coin archives. The most common one is the Query-By-Example (QBE) approach involving standard coin categories (mainly the coin creation period). But other query methods can be used as well.

One of the most appreciated is the one based on a top down approach that typically allows complex query building through a series of selections, commencing usually from large categories (Roman Coins, Greek Coins) and then going through subcategories (iconography, mints, physical characteristics, etc.) to reach the desired set of coins. Other useful interface should implement:

- Identification search, that allows the identification of a coin according to specific fields, such as the obverse design, the reverse design, the inscription, the metal and the diameter of the coin.

- Iconographic search, used to identify coins using the types of imagery found on coins (objects, animals, architecture etc.)

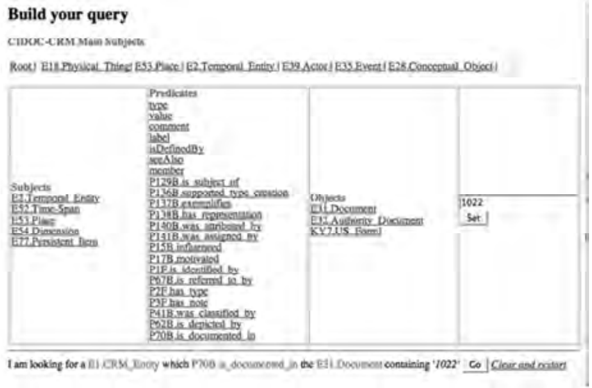


Figure 7. The Semantic Query Interface

All these interfaces are included in the COINS Management Tool but we have also added a semantic query framework to the preceding ones, using the integrated semantic information coming from coin archives and from the harvested information on images.

3.2.3 MAD: The Semantic Web Browser plug-in

The richness of the RDF graph model in which data are distributed makes it often difficult for users to get effective and meaningful data retrieval when only a simple or complex query interface is used, particularly when the graph grows in dimensions and complexity.

Sometimes it would be simpler and faster to browse the multidimensional graph structure allowing users to choose a starting point and move along different paths through the graph to reach the desired data. To allow this kind of data navigation we have developed a Mozilla Firefox plug-in based on SIMILE technology (Simile, 2008). The plug-in turns the Firefox browser into a semantic browser able to open, search and save RDF datasets. Browsing is based on the faceted browsing UI paradigm.

A facet in this view is a particular metadata element considered important for the dataset browsed. Users can select a starting point which they consider relevant for their search. The browser extracts a list of facets, their values, and the number of times each facet value occurs in the dataset. Then it's possible to add or remove restrictions in order to focus on more specific or more general slices of the model. A "free text" restriction can be also added to reduce the browsed dataset to all items that contain the required string . The interface was also configured to interact with the MAD container: all the semantic information stored therein can be browsed in order to retrieve relevant semantic objects and references to external resources. RDF resources on the Web can also be visualized by providing their URLs and saving them in the MAD Semantic Database (Figure 8).

Figure 8. MAD Semantic Browser Interface

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. We are also using MAD with the AMA toolset to create an on-line application for the complete semantic management of coin collections for the COINS Project. All this information is stored in RDF format and is ready to be queried, integrated and shared in the Semantic Web framework.

3.3 Interoperability

The mapping tool is based on AMA. It is able to perform the automatic mapping of metadata extracted from a site to RDF information encoded using the CIDOC-CRM ontology. While the crawler harvests the pages looking for relevant coins images, metadata coming from the harvesting process concerning the found images (stored in the spider's index), is processed and converted. A set of URI is also created to link our semantic metadata to the retrieved resources and the downloaded images to establish a bidirectional link between them.

The whole process is driven by a XSL stylesheet that we have created to map the metadata used by eBay pages to the CIDOC-CRM entities. This stylesheet is invoked by the AMA mapping engine (that is physically in charge for all the operations) for the on-the-fly conversion process, for the creation of the RDF datasets and for the definition of the URI referencing the images.

The semantic nature of this information needs a proper container to be stored and managed and to allow users from law enforcement or numismatists to build complex, articulated and diverse semantic queries in order to discover stolen coins according to semantic relations. It would be possible, for instance, to find all the coins similar to a given one but discovered in a certain place, coined in a given city, put online by an eBay user and bought by another one (a very complex query involving many CIDOC entities like places and people, but very easy to accomplish on a set of semantic-ready information). Another advantage given by this kind of approach is the possibility to combine image recognition and semantic information to build custom queries.

4. RESULTS

4.1 COINS spider

For testing purposes, the spider is running on a PowerMac G5 with 1GB of ram and a 2 Mbit ADSL connection. Threaded crawling requests to the eBay site are low in number, for politeness. After initial experimentations with URLs and depth of search (Section 3.2.3) crawling for images commenced with a small number of top N to test the image plugin algorithm. Having verified that the plugin works in a robust way, the topN limitation was removed.

Through a period of a week, the US site was slowly crawled in order to fetch coins links and images. The initial results are encouraging. 13450 urls were fetched from eBay US. 10906 of them (81%) are actual ebay items (cgi.ebay.com links). 2544 of those links were identified as potential coins candidates – after pruning of duplicate URLS the number was reduced to almost half. The image spider also fetched 1279 prospective coin images related to the candicate coins URLs. Out of those images, 180 (14%) are not actual coins but pictures of coin books, busts of Caesar, etc (Figure 9). Checking, parsing and saving of the images was complete in two hours.

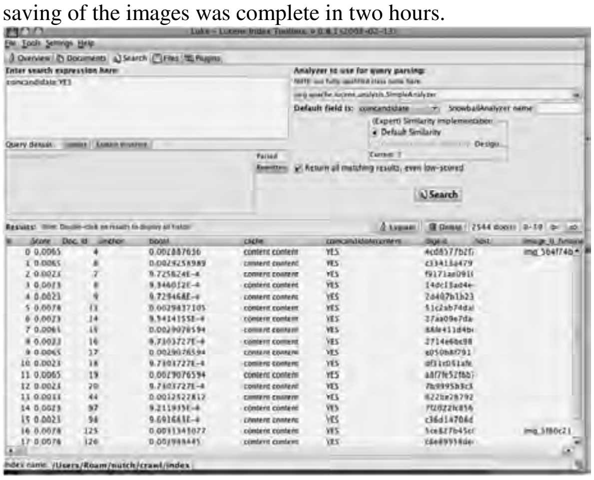


Figure 9: Index analysis

We consider these results very encouraging since we have not utilised any image recognition techniques to reduce the number of non-coins images. Numerous of the fetched URLs were from completely different parts of the eBay site but due to the profile screening they were correctly not identified as coins candidates.

4.2 Semantic database and connectivity

We have tested the semantic framework on the information harvested by the spider and the results seem promising. The mapping tool is affordable and fast and the RDF dataset generated by the mapping tool is formally and semantically correct. The rational structure of RDF triples seems to be very clean and is usually well rendered by the semantic. The URI's connections among RDF and images work in most of the cases. Some aspects of data encoding process, however, need to be refined and semantic information sometimes need to be manually integrated by the users.

5. CONCLUSION AND FUTURE WORK

The COINS project aims to assist in the fight against illegal trade of antiquities, and specifically ancient coins. This paper has examined initial research results achieved as part of the project in two distinct areas: the coins image search spider, and the semantic database. The spider is an open source implementation of a search engine which crawls the web for candidate coin images and related content. Its effectiveness was tested on the US site of eBay with encouraging results. In future work we wish to include more auction sites such as vcoin.com – screening of results will be achieved (apart from URL restrictions) by profile settings as was done for eBay. Additionally, we wish to examine further ways to lower the amount of false positives retrieved by the engine.

As far as the semantic database is concerned, for the future we are planning a refinement of the XSL stylesheet used by the mapping tool for the on-the-fly conversion. We aim to make it more powerful and to reduce the human intervention on CIDOC-CRM information store in the semantic container.
We would like to provide the possibility for users to manually add relevant semantic information to images including a tool like Annozilla (Annonzilla, 2008) to assist users in the creation of RDF annotation associated with the crawled web pages and to allow them to modify existing ones. This will enhance the level of detail in images description and will assist final users in building more powerful and complex semantic queries.

REFERENCES

AMA. 2008. *Archive Mapper for Archaeology*. Retrieved June 15, 2008, from http://www.epoch.eu/AMA

Annonzilla. 2008. *The Annozilla Project*. Retrieved June 05, 2008, from http://annozilla.mozdev.org/

Baluja, S. 2008. *Building software tools to find child victims*. Retrieved June 15, 2008, from http://googleblog.blogspot.com/2008/04/building-software-tools-to-find-child.html

Baluja, S., & Jing, Y. 2008. PageRank for Product Image Search. *World Wide Web Conference 2008* (pp. 307-315). Beijing, China.: ACM.

Benedict, L. 2004. *Comparison of Nutch and Google Search Engine Implementations on the Oregon State University Website.* Retrieved June 13, 2008, from Oregon University: http://www.misterbot.fr/OSU_Queries.pdf

Cafarella, M., & Cutting, D. 2004. Building Nutch:Open. Source Search. *ACM Queue , 2* (2), pp. 54-61.

CheckMend. 2008. *Checkmend press release*. Retrieved June 14, 2008, from http://www.checkmend.co.uk/pdf/CheckmendPressRelease20080518.pdf

Conklin, J. E. 1994. *Art Crime Book.* Norman, G. Great sale of the century. Independent, 24 and Art Crime Book by John E. Conklin; Praeger Publishers, 1994. 326 pgs. Praeger Publishers.

Croft, S. N., Doerr, M., Gill, T., & Stead, S. (2008). *Definition of the CIDOC Conceptual Reference Model.* Retrieved June 14, 2008, from http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.doc

Elkin, N. T. (2007). *Why coins matter: Trafficking in undocumented and illegally exported ancient coins in the North American marketplace.* Retrieved June 15, 2008, from SAFE features: http://www.savingantiquities.org/feature_coins.php

EPOCH. (2007). *Arrigo in Fedora: EPOCH project common infrastructure tool chain demo*. Retrieved June 15, 2008, from http://partners.epoch--net.org/common_infrastructure/wiki/index.php/Arrigo_in_Fedora

EPOCH. (2008, May). *EPOCH European Network of Excellence in Open Cultural Heritage*. Retrieved June 14, 2008, from http://www.epoch.eu

Felicetti, A. 2006. MAD: Managing Archaeological Data. In M. A. Ioannides (Ed.), *The e-volution of Information Communication Technology in Cultural Heritage. The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage.* Nicosia, Cyprus: VAST.

Felicetti, A., & Lorenzini, M. 2007. Open Source and Open Standards for using integrated geographic data on the web. In D. Arnold, A. Chalmers, & F. Niccolucci (Ed.), *Future Technologies to Empower Heritage Professionals. The 8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage.* Brighton, UK: VAST.

Felicetti, A., & Mara, H. 2008. Semantic Maps and Digital Islands: Semantic Web technologies for the future of Cultural Heritage Digital Libraries. First Workshop on Semantic Interoperability in the European Digital Library. *First Workshop on Semantic Interoperability in the European Digital Library.* Tenerife, Spain.

Guha, R., McCool, R., & Miller, E. 2003. Semantic search. *International World Wide Web Conference* (pp. 700-709). Budapest, Hungary: ACM.

Nutch. 2008. *Nutch search engine homepage*. Retrieved June 15, 2008, from http://lucene.apache.org/nutch

Simile. 2008. *The SIMILE Project*. Retrieved June 18, 2008, from http://simile.mit.edu/

SPARQL. 2008. *SPARQL Query Language for RDL*. Retrieved June 13 2008, from http://www.w3.org/TR/rdf-sparql-query

Todeschini, P. W. 2007. the Medici Conspiracy: The Illicit Journey of Looted Antiquities from Italy's Tomb Raiders to the World's Greatest Museums. *British Journal of Criminology , 47* (2), 367-370.

W3C. 2008. *Home of the Semantic Web Initiative*. Retrieved June 14, 2008, from http://www.w3.org/2001/sw/

Wutkowski, K. 2007. *U.S. retailers want online sellers to fight theft*. Retrieved June 14, 2008, from http://www.reuters.com/article/technologyNews/idUSN2560973520071025

Zhang, Z., Rojas, C., Nasraoui, O., & Frigui, H. 2006.. SHOW AND TELL: A Seamlessly Integrated Tool For Searching with Image Content And Text. *ACM-SIGIR Open Source Information Retrieval workshop.* ACM.

SPECIFICATIONS AND STANDARDS FOR THE GEOMETRIC DOCUMENTATION OF CULTURAL HERITAGE

S. Tapinaki [a], A. Georgopoulos [b]

National Technical University of Athens, Laboratory of Photogrammetry
Iroon Polytechniou 9, GR-15780, Athens, Greece
[a] sevi@survey.ntua.gr
[b] drag@central.ntua.gr

KEY WORDS: Geometric Documentation, Cultural Heritage, Specifications, Standards, Database Management System

ABSTRACT:

Documentation of monuments is essential for the protection of Cultural Heritage. Unfortunately, there is lack of Specifications and Standards, worldwide. This paper describes an attempt to collect and evaluate international experience for this subject, towards the compilation of specifications, standards and guidelines especially for the Geometric Documentation of Greek Cultural Heritage. The central idea is their integration in a Data Base Management System for project administration, project supervision and data management. The system will provide the definitions of specifications for the products and deliverables, while guidelines will ensure that these specifications are met. The existence of modern and adaptable technical specifications will definitely facilitate the use of alternative solutions, for geometric documentation of monuments, even of large and complicated ones. Project managers will have the capability to supervise the application of different procedures in a certain project, while, at the same, since data will be stored constantly and continually in the system, monitoring of the monument will be possible. Additionally, quality standards concerning management of projects, spatial data, images, large scale drawings etc. are deliberated and evaluated, so that they can be integrated into the system, for the quality certification of the products. Such an action may lead to the compilation of ISO standards in the future, if such an action is considered appropriate.

1. INTRODUCTION

The protection of Cultural Heritage is a very complicated issue that requires inter-disciplinary approach and depends on related legislation, historical and constructional information and other related parameters. The objective of Geometric Documentation is the recording of the existing status of monuments and the production of vector drawings, orthophotographs and 3D digital models of the object. Monuments differ depending on the type, the era of construction, the architectural style, the history and the symbolism. All these factors, along with the available equipment and know-how, must be taken into consideration, during the documentation of a monument, in order to ensure that the products are complete and reliable.

In the survey of a monument difficulties are present, mainly caused by the nature of the monument itself: high relief, anaglyph, unknown surfaces and forms, different materials, damages, special architectural constructions are only but a few of the main sources of problems. Moreover, difficulties appear during field work: access, photographing, marking of control points etc. Development of technology has provided solutions in a lot of problems, which, in earlier times, could not be confronted, in order for the results of the geometric documentation to satisfy the requirements of accuracy and completeness of the final products.

In Greece, but also internationally, systematic specifications that would follow and exploit the technology development do not exist. For each case of monument, special specifications are compiled, so there is no cohesive confrontation of the problem and each time the results depend on the experience of whoever compiles the specifications and not on the objective needs for each case of Cultural Heritage object.

The survey of a monument may last from a few weeks to a couple of years, according to the size and complexity of the object and the demands of the project. Also, different and multiple methods are applied to the same project, such as survey measurements of control points, photographing of the object, mathematical calculations for the determination of control points coordinates, application of photogrammetric methods for the production of orthophotographs and other image based products, terrestrial laser scanning and post processing methods of point clouds for 3D digital models of the object, etc. Therefore, project administration and supervision are very difficult. Furthermore, the data collected and produced after processing are diverse and enormous in size, e.g. images, notes, field sketches, survey measurements, points, coordinates, vector drawings, combined drawings, orthophotomosaics and many more, so proper data management is a difficult task too (Tapinaki, 2005). Standardisation between projects is also a vital requirement, especially for organisations that deal with a large number of projects each year (Barber, 2003).

The results of Geometric Documentation are usually being used as background in studies of monument restoration, where high accuracy is usually required. Somehow quality and accomplishment of accuracy must be ensured. The nonexistence of specifications and standards creates, also, problems in the communication between those who carry out the survey, the administrators, the supervisors and finally -and most importantly perhaps-, the users.

It would be useful, if the results and the acquired knowledge and experience of a certain project could be spread and used worldwide, so that future surveys would benefit from previous errors. Thus, if a similar case comes up, knowledge from previous projects could be used and implemented immediately, without wasting time for solving the same problems anew. This would be achieved only by the standardisation of procedures and products.

Given the diversity and the complexity of Greek monuments (historic buildings, structures, statuary, objects) and archaeological sites, there is a strong need for the compilation of Specifications and Standards that will take into consideration and will expose the rules, during a project, from the initial decision making phase, until the final delivery. In Figure 1 a complex procedure for an integrated approach of the Geometric Documentation of Monuments is presented.

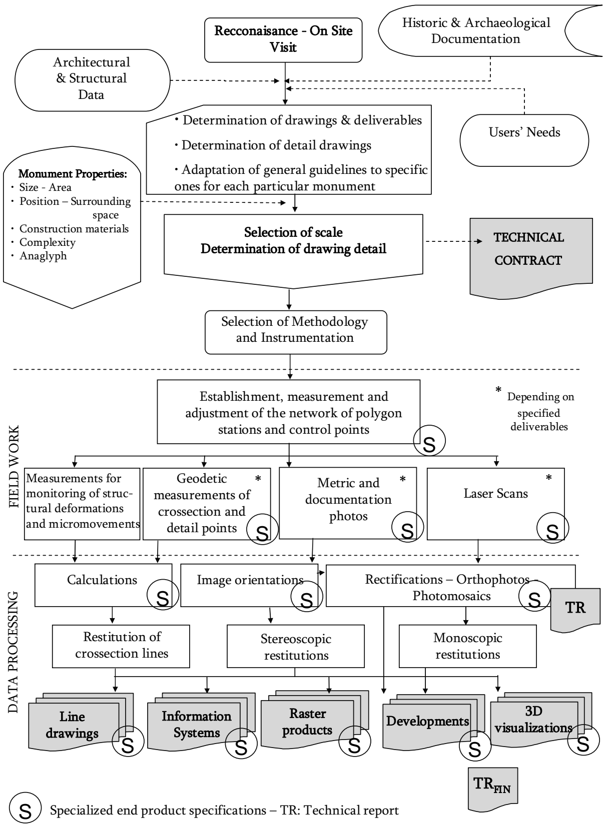


Figure 1: An integrated approach for the Geometric Documentation of a Monument

A Data Base Management System, for project administration, project supervision and data management that will provide guidelines, standards and specifications for geometric documentation of monuments, is also considered to be a very useful tool.

2. THE DATA BASE MANAGEMENT SYSTEM

2.1 The DBMS Architecture

The main scope of the DBMS is to facilitate administration, supervision and data management of geometric documentation projects, while, at the same time, it will provide guidelines and help with the compilation of standards and specifications.

The design of the DBMS is, at the time, at a preliminary stage. General concepts for the architecture are under consideration, while accumulation and analysis of user demands are in progress. A schematic representation of the system is shown in Figure 2.

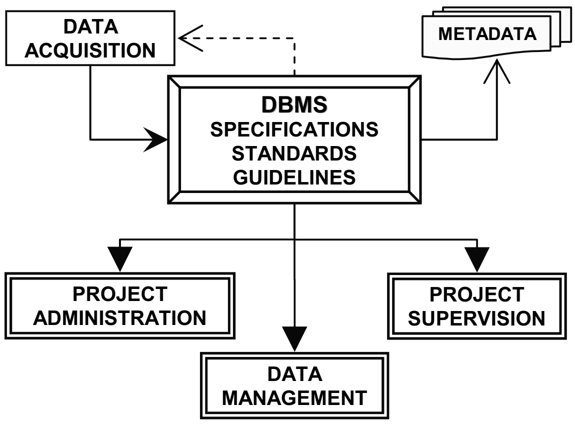


Figure 2: The DBMS Architecture

The system should provide the necessary tools and guidelines for:

- Data Acquisition
- Data Entry
- Data Management
- Metadata Management and extraction
- Project Administration and Supervision

For the facilitation of the design of the whole system, it is divided into subsystems. The subsystems are being designed separately and will be integrated together when ready and functional.

The system should integrate and provide the definitions of specifications and standards for the products and deliverables. Certain parameters will be changeable, so desirable and allowable choices could be made, according to the requirements of a specific project. At the same time, additions or deletions of other parameters will be possible. In such a way, specifications and standards could be adapted in the best possible way to any case of monument.

Guidelines, on the other hand, will suggest directives, according to the desirable results, to ensure that the specifications and standards are met. They will be organised in such a way that guidance will follow the work to be done, which results to reduction of mistakes and gain of time. To mention but a few guidelines for the programming of field and office work will direct the administrators during decision making, standard forms in text format will be provided for data acquisition, directives for the preparation and presentation of final results will be given etc.

The system will provide two main functions: management and retrieval. In this way, the flexibility of the system and the safety of data are ensured, as different access rights can be defined for various groups of users. The management function provides operations for addition, abstraction or modification of data, while the retrieval allows only search and retrieving operations.

The determination of the definitions of all the concepts and relationships used in geometric documentation of monuments is essential. The development of an ontology that will fulfill the system and ensure the demanded interoperability is currently under consideration. For this scope, various alternative and general approaches are under evaluation, with the CIDOC CRM to be the most promising (http://cidoc.ics.forth.gr/). 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. The CIDOC CRM is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework that any cultural heritage information can be mapped to. It is intended to be a common language for domain experts and implementers to formulate requirements for information systems and to serve as a guide for good practice of conceptual modelling. In this way, it can provide the "semantic glue" needed to mediate between different sources of cultural heritage information, such as that published by museums, libraries and archives. Since 9/12/2006 it is official standard ISO 21127:2006" (CIDOC CRM). This semantic model is referring to the documentation of cultural heritage in general and it could be used as a base for the ontology development especially for geometric documentation of cultural heritage, with the appropriate adjustments in order to be supplemented with the necessary concepts and relationships.

2.2 Project Administration and Supervision

For the design and completion of the subsystem for project administration and supervision, all the needs of administrators and supervisors are being gathered and analysed. A detailed questionnaire is being compiled to ensure that all aspects will be covered. The questionnaire is intended for all experts involved in monument geometric documentation projects and also for those who use the results of such projects.

Certain functions of the system will facilitate project administration by guiding the administrators through every phase of the project, while at the same time functions for continuous project supervision will be provided.

For the compilation of the contract, there will be a standard one with guidelines, which with the desirable changes (deletions or additions) could be adapted on the certain project. Further, cost of the project will be estimated, by following certain guidelines. During the project, inspection of work will be possible, at any time. Administrators could edit the requirements and update the system with possible necessary changes, for example the system must be continually updated for the composition of personnel. Standard technical reports (special contract compilation, Project form, periodic and final reports) and guidelines for deliverables will be provided by the system, so administrators could give the proper instruction for their preparation.

2.3 Data Management

The subsystem for data management is based on the DB "Geometric Documentation Data", which was designed especially for this purpose (Tapinaki, 2005). The Database is able to support the data management for geometric documentation of a monument during all its phases. It was initially designed to cover mainly topographic and photogrammetric methods, so readjustments have to be done, towards its completion with terrestrial laser scanning methods. Also, its conceptual modeling was made using the Entity - Relationship model (E-R model) and the logical modeling with the Relational model. Redesigning the system with other models, like object-oriented and semantic ones, would perhaps make the system more stable, powerful and capable to manage more diverse data.

The goal is the management of the data to follow the course of fieldwork. In this way, the data collected or produced will be immediately entered and updated, so it may be easily restored at anytime. The necessary forms may be extracted to facilitate the collection of data. These forms will be blank and ready to be used for the recording of data acquired in the field. In this way, the completeness of data acquisition is ensured. For example, during image acquisition certain parameters should be recorded, which are vital for image post processing.

When data is stored in the database, it is being codified with a unique alphanumeric code. In this way the immediate identification of its attributes is possible, since each digit of this code represents the value of a certain attribute. Additionally, codification of data is vital both for data acquisition and post processing. For example:

(a) Each projection plane is being coded with a 4digit code

1 2 3 4

where:

- 1: represents the kind of projection. This digit may take the values "1", "2" or "3" for the three kinds of projection plane: façade, horizontal section or vertical section, respectively.
- 2: represents the type of projection. The value for this digit is "0" in case of a façade; "1" or "2" to represent either floor plan or ceiling plan in case of a horizontal section; or "1" or "2" again to denote the parallelism or perpendicularity to the main axis of the object in the case of a vertical section.
- 3: represents the serial number of the projection plane
- 4: represents the orientation of the projection plane, i.e. where it is facing (S=South, N=North, E=East, W=West, U=vertically Upwards, D=vertically Downwards)

All data referring to a particular projection will have the appropriate reference into their code.

(b) Each image is being codified with a 10digit number

123 4 5 6 789 10

where:

- 123 4: is the 4digit code of projection plane as it is explained above.
- 5: the code of the camera that is used for the image acquisition. This code is user defined.
- 6: the type of image post processing (e.g. R=Rectification or O=Orthophotography)
- 789: image serial number
- 10: the type of film in case of analogue camera (C=colored or B=black and white), otherwise a single "d" represents digital image

Some examples for the codification of projection planes are shown in Figure 3.

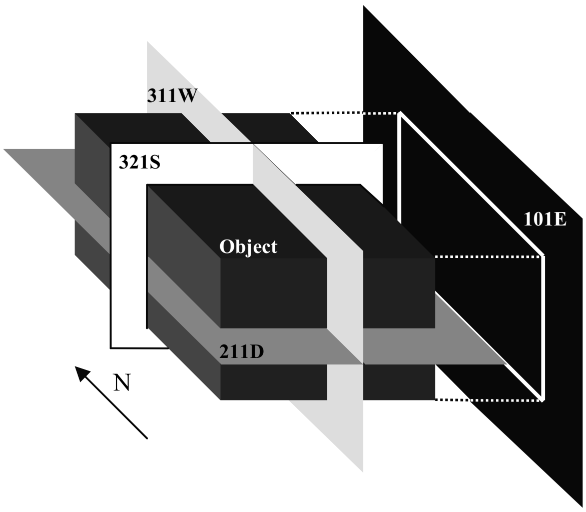


Figure 3: Examples of Coded Projection Planes

2.4 The Potential of the System

Such a DBMS is capable of covering all kind of monuments, from the most complex to the simplest ones. According to the needs of the final user in each case, the system will organize geometric documentation projects by standardising products and procedures, towards the safeguard of final quality and interoperability, while in addition it will facilitate certain procedures:

Project Administrators will have a very useful tool for complete management of projects, from the initial programming of work till the final delivery. Timely intervention in eventual problems, correction of errors and covering of omissions will be possible. Also, they will have the capability to supervise the application of different procedures in a certain project. Also, since data will be stored constantly and continually in the system, there will be the possibility for monitoring of the monument during the course of works, but also this may continue afterwards.

Project supervisors from employer and contractor will be able to inspect, whenever they want, the progress of work and make the appropriate remarks and readjustments, always according to the agreed contract and specifications.

After insertion and update of data, management and usage of *metadata*, at any stage, will also be possible.

Significant advantage of the system is the extraction of *reports*, which help data acquisition, facilitate the insertion in the database and, moreover, they can be used in the compilation of technical reports, which accompany the delivery of final products.

Finally, the communication between those who carry out the geometric documentation, the administrators, the supervisors and the final users will be greatly facilitated.

3. SPECIFICATIONS AND STANDARDS

It is confirmed that specifications and standards for geometric documentation are essential for the interoperability, worldwide usage and comprehension of the results (Ioannides et al., 2005). During the last years, a lot of organisations and individual researchers are dealing with this subject; and most of them have already reached a very satisfactory point.

"The International Council on Monuments and Sites (ICOMOS), the Getty Conservation Institute (GCI) and CIPA together formed the RecorDIM (for Heritage Recording, Documentation and Information Management) Initiative partnership. The purpose of the initiative (started in 2002 end closed on 2007) was to bring information users and providers together to identify the nature of the gaps between them, to develop strategies to close the gaps and to recommend a framework for action." The Task Group 16 within this initiative has compiled a draft report on International Heritage Documentation Standards with recommendations for work practice, technical specifications and data standards (RecorDIM TG16, 2007).

3.1 Specifications for Greek Cultural Heritage

For the determination of the Geometric Documentation type that will be applied, certain questions must be answered first:

- What is important for the specific project?
- What is the distinctiveness of the specific object?
- What is the required level of detail in the final products?
- What is the future purpose and usage of the results?

It is obvious that, in most cases, the required results determine the methods. Therefore, the basic principle, in the compilation of specifications that will regulate the implementation of geometric documentation of Greek monuments, must be the clear description and the objective control of products and not the analysis of the methods that will be applied for their production. The needs, of those who use the results of Geometric Documentation for other studies, must be accumulated and analyzed. It is a case of specifying the end result and not methods to be employed.

For the compilation of these specifications, existing material, as Greek legislation and decisions of Greek competent organisations that deal with the restoration or maintenance of monuments, is being studied (Hellenic Ministry of Culture, Technical Chamber of Greece, etc.). Also, the results of research of International Organisations are evaluated, towards their adjustment in Greek Cultural Heritage (RecorDIM, National Park Service USA, English Heritage, Object id etc.). The main goal is the revision, the enrichment, the filling and sometimes the compilation from the beginning of this material, in order to adjust it and adapt it to the specific Greek needs. The intention is to completely cover each object and the provision of full exploitation of up-to-date technology. Thereby, quality and economic implementation of the project will be ensured.

The specifications should be clear, complete and most of all flexible, so that with the omission of certain paragraphs or the addition of specific requirements, they may be adapted to the special needs of each object. Consequently, they may be used as they are or as the base for the compilation of special specifications for a certain work. There will be an effort that these specifications will give the possibility of application of any topographic, photogrammetric or other method, depending on the object and utilisation of up-to-date equipment, in order to improve the quality of products and the reduction of cost.

3.2 Contents of Specifications and Standards

Administrative rules will be imposed for the assignment, observation, supervision, delivery and acceptance of the project, along with technical rules for the distinctly clarification of type, accuracy and presentation of the results.

The specifications will comprise, basically, the following definitions:

- A *Glossary* with the significances of the terms being used
- The *Policies* for compilation and application of the project, according to International Heritage Charters and Declarations
- Determination of the *Types* of Geometric Documentation, in proportion to the purpose of the project, usage of products and required level of detail
- Standards and requirements for *Contract compilation, Timetable, Deliveries and Estimate* (cost).
- Policies for *Project Administration* and *Supervision*
- *Safety Demands* and *Requirements*. Danger evaluation and safety measures should be stated, concerning the monument itself, equipment, data and personnel. Insurance of equipment along with office and field staff must be taken into consideration.
- *Data security,* data management and archiving
- The composition of *Personnel* and the conditions of their participation to the project.
- Standards and requirements for the *Equipment*, the most significant being the of calibration reports.
- *Needs* and *Requirements* of the employer, these have to be determined and clearly mentioned in the contract, in order to avoid misunderstandings.
- *Obligations* of both the employer and contractor
- Determination of the *evaluation* and *control* of the final products for accuracy, completeness and closeness, in order to ensure their sufficiency and suitability.
- Determination of the handling and protection of *Copyright.*

All the above mentioned issues along with others, minor ones are under consideration and testing for the completeness of specifications and standards.

4. GUIDELINES

To ensure the accurate implementation of specifications, and the meeting of standards, guidelines are being composed. These guidelines refer to the application of work and procedures, according to the required accuracy. Guidelines presented from international organizations are being thoroughly studied, for their adjustment to the Greek needs. Guidelines are being composed simultaneously with specifications and standards.

Directives will be given, in all phases of the project, concerning:

- The compilation of *Contract,* determination of *Timetable* and *Deliveries* and estimation of *Project Cost*
- *Programming of Work*: Guidelines should be provided for all parameters that have to be considered and determined for programming of field work (image acquisition, topographic measurements etc), as well as work in the office.
- *Data Acquisition*: Sketches, Topographic measurements, image acquisition, notes etc. Sketches of the object, where various details are recorded, during field work, have to be clear and specific, because they are being used during data post processing. It is important that specific guidelines will give directives for the compilation of sketches.
- *Data management* and *archiving* (Data Codification, materials, media etc)
- *Preparing* and *Presentation* of *drawings* and *products.* For the production of vector drawings various layers with diverse data have to be managed. Guidelines for names, colours, line weights etc. would facilitate this procedure, and standardise the resulting digital drawings. Also, a standard drawing template should be provided.
- Compilation of *Technical Reports* (special contract compilation, Project form, periodic and final reports)
- Rules for *Accuracy, Completeness* and *Closeness evaluation* and *control*

In addition, several other topics are under discussion on which guidelines should refer, like permanent or temporary marking of control points, data and metadata post-processing etc.

5. QUALITY STANDARDS

Quality management and certification is very important, since, geometric documentation of monuments is required, usually, for restoration or maintenance projects, where high accuracy and quality is demanded.

Quality Standards concerning management of projects, spatial data, images, large scale drawings etc. are deliberated and evaluated, so that they can be integrated into the system, for the quality certification of the products (e.g. ISO, ANSI etc.). Quality standards will, basically, affect the guidelines, by applying the appropriate quality controls whenever it is necessary. The objective is that during implementation of certain procedures, qualitative monitoring will be applied; in order for the final results to conform to the quality specifications, both as far as accuracy and presentation are concerned.

6. CONCLUSIONS AND FUTURE OUTLOOK

The existence of specifications, standards and guidelines is an international issue, the usefulness and necessity of them is still under question. This is an attempt (a) to demonstrate that the lack of specifications creates a lot of problems, (b) to introduce the compilation of specifications, standards and guidelines especially for the Geometric Documentation of Greek Cultural Heritage and (c) to present the development of a Data Base Management System for project administration, project supervision and data management of geometric documentation projects. The latter task is the most difficult and seems to be very complicated, but the partition of the system into several subsystems facilitates its understanding and design, considering that the design of the subsystem for data management, which is the most complex, is already at the final phase.

The requirements of such a system may increase work, since more information should be recorded and continuous update is necessary. On the other hand, certain procedures will be faster so the cost of project will be positively affected. Finally, quality, interoperability, security and safety, free access and usage of data will be ensured. Moreover, repetitions in future studies of the same object are going to be avoided.

Communication between those who carry out the geometric documentation, the administrators, the supervisors and the final users will be facilitated. Last but not least, the usage of specifications will allow the diffusion and spreading of knowledge.

REFERENCES

Bryan P., Blake B., 2000. Metric Survey Specifications for English Heritage. English Heritage, York and Swidon. 111 pages

Hellenic Ministry of Culture, 1995. General Specifications for Byzantine and Post-Byzantine Monuments Studies. Directorate of Byzantine and Post Byzantine Monuments Restoration, Greece. 90 pages

Ioannides M., Georgopoulos A., Scherer M., 2005. Standards in Cultural Heritage: The Missing Grammar for the Digital Documentation of the Past. *Proceedings of XX CIPA International Symposium*, Torino, Italy. 861-870

Tapinaki, S., Georgopoulos, A., Sellis, T., 2005. Design of a Database System for Geometric Documentation. *Proceedings of XX CIPA International Symposium*, Torino, Italy. 969-973

Technical Chamber of Greece, 1992. Specifications and Cost of Architectural Documentation Studies for Restoration of Monuments. Technical Chamber of Greece, No 1694, 07/01/1992, Greece.

References from websites:
Barber D., Mills J., Bryan P., 2003. Towards a Standard Specification for Terrestrial Laser Scanning of Cultural Heritage. *Proceedings of CIPA XIX International Symposium.* Antalya, Turkey. http://cipa.icomos.org/index.php?id=357 (last accessed 18 June 2008)

CIDOC CRM. http://cidoc.ics.forth.gr/ (last accessed 18 June 2008)

CIPA. http://cipa.icomos.org/index.php?id=362&PHPSESSID=36bc0a617f19d67e8cf2f465e6f237b9 (last accessed 18 June 2008)

English Heritage. http://www.english-heritage.org.uk/ (last accessed 18 June 2008)

National Park Service. http://www.nps.gov/history/hdp/standards/index.htm (last accessed 18 June 2008)

Object id. http://www.object-id.com/ (last accessed 18 June 2008)

RecorDIM TG16, 2007. Draft Report on International Heritage Documentation Standards. http://cipa.icomos.org/index.php?id=395 (last accessed 18 June 2008)

RecorDIM. http://extranet.getty.edu/gci/recordim/index.html (last accessed 18 June 2008)

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

METADATA-BASED TERMINOLOGY ONTOLOGY FOR KNOWLEDGE MANAGEMENT OF AN ARCHITECTURAL HERITAGE IN DANGER

E. Andaroodi [a,*] , K. Ono [a], A. Kitamoto [a]

[a] Research Organization of Information and Systems, National Institutes of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo, Japan - (elham, ono, kitamoto)@nii.ac.jp

KEY WORDS: Ontology Knowledge Model, Building Typology, Terminology, Object ID Metadata Standard

ABSTRACT:

A multilingual and lexical schema as part of a metadata-based ontology knowledge model for the domain of architecture, specifically for the building typology of a UNESCO world heritage site in danger, is introduced in this paper. The discussion starts with the related works on the domain ontologies for cultural heritage. The architecture of our ontology consisting of two complementary multilingual-lexical and Object ID metadata schemas is presented as a set of interconnecting instances linked with the properties of taxonomy of classes. The discussion continues with the implementation of the ontology schema inside the knowledge representation tool and instantiation of the terminology for building types. It ends with an assessment of the consistency of the ontology from the technical point of view and a content-based evaluation. The paper concludes by proposing the application of this ontology for semantic knowledge management of multimedia content for the selected case study over the Internet as well as a discussion on future work.

1. INTRODUCTION

1.1 Cultural Heritage Documentation

Recognition of cultural heritage, which can be considered as the initial process of conservation starts by documentation of an historical relic. Advanced documentation deals with the management of data of architectural buildings at different levels by using advanced knowledge management techniques. It provides raw preliminary data, classified data as information, and systematically analyzed and structured data as the knowledge of the domain. Its real importance is revealed in its capacity to keep an accurate record of buildings for numerous relics and limited conservation resources.

1.2 Knowledge Management

Knowledge management can provide a powerful framework for archiving, semantic annotation, and the distribution of visual or textual data of cultural heritages, especially over the Internet, for access by an unlimited number of users. Such wide network access to numerous databases creates new complexities and challenges for retrieving desired data.

Just as a simple example, a recent keyword search using the terms “cultural heritage” provided 17,500,000 web page hits. How do these web pages contribute to the documentation and correct recognition of cultural heritages? How well do they provide access to their information? How can knowledge management systems help to solve these problems? The design and application of a knowledge model for each homogeneous field is a key issue to conceptualizing the content, specifying it, and representing its technical knowledge in such a way that it can be machine understandable. Such models need to follow sets of standards for the data exchange over networks for better human and machine communications.

1.3 Ontology Knowledge Model

The study of ontology design and development is key solution for knowledge management. Ontology can provide “specifications of a conceptualization” (Gruber, 1995) and the conceptualization of a domain into computer understandable languages, such as the Resource Description Framework or Ontology Web Language. It gives structure to data through the knowledge acquisition techniques used in object oriented knowledge representation tools. Ontology is a possible solution for providing a more sophisticated search and retrieval for the structured information in a more advanced generation of World Wide Web, the semantic web (Berners Lee et al., 2001).

In this paper we address the design process of ontology-based knowledge management for the historical architecture domain for a case study on a site within UNESCO World Heritage in danger, the Citadel of Bam. Our ontology conceptualizes the typology of the buildings of the target site under two complementary schemas. At first the lexical and multilingual taxonomies (Andaroodi, 2006) and then, the Object ID metadata (Thornes et al., 1999) that is merged and revised in the target ontology.

The rest of the paper is structured as follows. Part 2 discusses the architecture of the lexical schema and the typology terminology of the Citadel of Bam. Part 3 explains the multilingual schema and the hierarchy of thesaurus for each type of building that includes a discussion on the challenges of a multilingual terminology. The implementation of the ontology schema by using the Protégé knowledge representation tool is also explained in this part. Part 4 discusses the evaluation and assessment of the ontology and introduces the proposed application for the semantic knowledge management of the multimedia content of a case study over the Internet. Part 5 concludes.

2. ONTOLOGY FOR BUILDING TYPOLOGY

Ontology development provides a taxonomic schema through the design of hierarchies of classes, the information related to classes as properties, and the different types of data given in properties as instances or attributes of each class (Noy et al., 2000). The domain's data is represented in the ontology as a taxonomic lexical specification of the components with their relationships through an elaborate analysis and synthesis that will also help to recognize the target domain.

Ontology knowledge management can provide well structured schema especially suitable for large corpora. We selected a historical city and citadel that was destroyed in an earthquake in 2003 as part of Bam and its cultural landscape, as our case study. The site comprises several types of buildings with complicated semantics and evolutions over time. This provided us with a valuable resource of information that needed to be well organized through our ontology knowledge model.

2.1 Related Work

Ontologies are developed and designed for a wide range of different types and disciplines. General Terminology Ontologies, such as the United Nations standard products and services code, UNSPSC (http://www.unspsc.org/), Word Net, an online lexical database for the English language (http://wordnet.princeton.edu/perl/webwn), or upper level ontologies such as SUMO (http://www.ontologyportal.org/), which is the standard upper level ontology for general concepts, or domain ontologies such as Geography Ontology, can be mentioned. Here we will present state of the art specific ontologies developed for the management of cultural heritages. A brief survey of the ontology-based metadata is also presented. Ontologies for specific domains conclude the review. It must be mentioned that we were faced with a lack of architectural domain specific ontologies among the related works.

2.1.1 Art and Architecture Thesaurus: The AAT is a thesaurus, a structured vocabulary that can be used to improve the access to information about art, architecture, and material culture. It contains around 125,000 terms and other information about the concepts. "The terms in AAT may be used to describe art, architecture, decorative arts, material culture, and archival materials in a taxonomic environment. The coverage of the AAT ranges from antiquity to the present and the scope is global". AAT also provides structure and classification schemes that can aid in documentation (Peterson 1994).

2.1.2 Reference Ontology for the Interchange of Cultural Heritage Information (ICOM CIDOC): "ISO 21127, ICOM CIDOC is domain ontology for cultural heritage information. As such, it is designed to be explanatory and extensible rather than prescriptive and restrictive."

The ICOM CIDOC ontology tries to cover different fields of cultural heritage information; therefore, the instances are more generic with a wide area proper to the exchange between libraries, museums, and archives.

If the target of the ontology is a specific narrow domain of architectural relics, it cannot cover the desired concepts or relationships. For this purpose, especially for the metadata management of historical buildings, object ID is more properly suited.

2.1.3 Object ID Ontology: "Object ID is an international standard for describing cultural objects as a metadata standard containing specific items. It has been developed through the collaboration of the museum community, police and customs agencies, the art trade, insurance industry, and values of art and antiques" (Thornes et al., 1999).

For this research, the object ID metadata schema has mapped with multilingual schema. It was implemented for the metadata gathering and knowledge input of around 50 types of buildings as the instances in the lexical schema.

There are some specific features proper for the metadata of Middle Eastern cities made from mud that is not covered by the schema of object ID. It is more a common list of attributes and it must be optimized for specific styles of buildings.

2.1.4 Ontology for Antique Furniture: Wielinga et al. (2001) explained about an ontology for the subset of art-object descriptions of antique furniture. The authors discussed the problem of acquiring background knowledge for the indexing and retrieval of images using ontology and semantic descriptions. It addresses the designing ontology for antique furniture.

2.1.5 Ontology for Archaeological Objects: Soo et al., (2002) developed a sharable domain ontology and thesaurus to help with the retrieval of historical images of the first emperor of China's terracotta warriors and horses. The authors described the process of designing the ontology as outlined below:

- Design tools to annotate historical images using a domain specific thesaurus and a sharable ontology;
- Design a query parser;
- Match the XML query schemas with the annotated schemas of the images in an RDF instance and find the most suitable image

2.1.6 Ontology for Images of Art: Smeulders et al (2002) discussed a research about the techniques for the enhanced access to cultural heritage data beyond a keyword search. The authors divided their approach into three main parts:

- User-tailored access to digital repositories;
- Ontology-based access to heterogeneous documentation (This part presents the interface of a tool for semantic annotation and a search for art images. The ontologies contained in the tool are taken from standards or thesauruses, such as AAT, Word Net, and Icon Class, and those presented in the RDF or RDFS classes. The tool provided an annotation template which was also taken from a metadata standard, the VRA 3.0 Core Categories).
- Access to the content of the image using invariant computer vision

2.2 Ontology Design Process

The process of ontology design starts with the study of terms and the lexical information related to it. In the case study for this research each term represents a type of building. Therefore, the term is linked with the name, type, and function of the building. The lexical schema is designed in different languages to provide a multilingual terminology for the typology of the Citadel of Bam.

Ontology can serve as an advanced lexical knowledge base in this stage that is similar to a dictionary or a thesaurus. The next important step is the specification, which is provided by the supporting schemas, that is given to the terms in the ontology. The structure of this schema at first presents the lexical specification of a term and then the characteristics of the term as an entity in the domain of historical architecture. The major background knowledge for our second supporting schema is the Object ID metadata standard. We have selected items from the Object ID list for cultural heritage (Thornes et al., 1998) that are proper for the Citadel of Bam, a Middle Eastern desert city made from mud brick.

Figure 1: Photographic view of citadel of Bam from top of city

2.2.1 A Case Study, Citadel of Bam: The case study for the building typology ontology is a historical citadel situated in the city of Bam. This site possibly originated in 500 B.C. and was used as part of the city until the end of 19th century. It had a wide variety of types of buildings that were situated in different districts (Figure 1). It was considered one of the biggest and oldest mud-brick cities that had more than 20 centuries of continuous habitation (Mehriar, 2003). This site is unfortunately seriously damaged after an earthquake in 2003 (Figure 2), and in 2004 "Bam and its Cultural Landscape" was registered on the list of world heritage in danger by UNESCO*.

Figure 2: General view from Bazaar after earthquake

Our ontology knowledge model for the typology of buildings in the Citadel of Bam is part of an upper level ontology for conceptualizing the process and outcome of the research project of a three-dimensional reconstitution and virtual reality demonstration of the Citadel of Bam (Ono et al., 2008). The part of the ontology for the typology discussed in this paper supports the semantic of 3D models of the buildings. Since different buildings in this huge site had complicated semantics, (function, components, and metadata) it was necessary to support a well-structured schema in order to represent the knowledge of each building in the site and connect it to the multimedia content as the output data for 3D models. The developed ontology contributed to the design of a knowledge-based website for the process and subsequent outcome of the 3DCG reconstruction of the site.

2.2.2 Architecture of Ontology: The architecture of the ontology follows the object-oriented structure (Rumbaugh et al., 1990) of the ontology editing tools. The pieces of information in such an environment are the entities with a specific identification number that are connected by links to properties or slots. A combination of IDs and properties belong to taxonomy of classes. Figure 3 shows the architecture of the target ontology of this research. This figure presents architecture of the ontology consisting of three different levels.

The first level specifies the lexical ontology. Lexical attributes such as a *word* or *semantic* are defined as the properties or slots belonging to the classes of the lexical ontology.

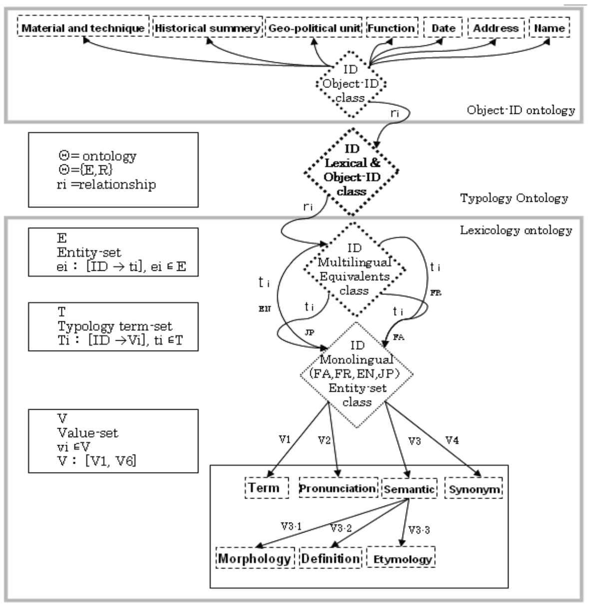


Figure 3: Architecture of building typology ontology consisting of two schemas as terminology and Object ID

Second level describes a multilingual ontology. The equivalent terms with lexical attributes used as the entities for each class of languages are linked by an ID that belongs to the property (or slot) *equivalent* to provide a multilingual and lexical entity-set for each type of building of the target corpus.

The upper part of the chart represents the Object ID ontology. The Object ID schema represents the items introduced in this metadata standard; it is simplified and merged into our ontology model. As can be seen metadata attributes (such as the *address, building material and techniques, dating, and physical condition*) are defined as classes. The instances of these classes

* http://whc.unesco.org/en/list/1208

are defined as IDs with related information given in the properties. For this purpose, the name and function of a building is defined in the lexical ontology schema.

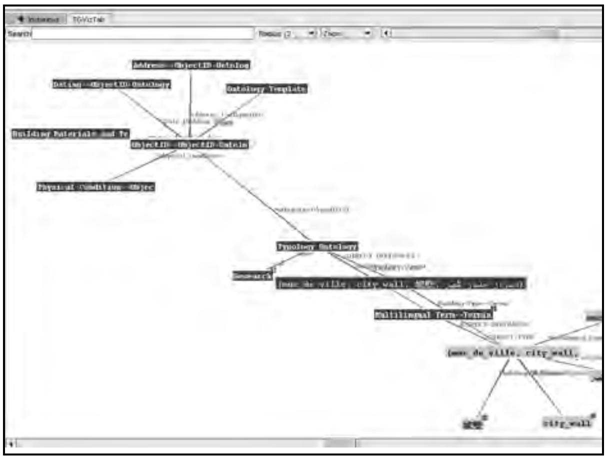


Figure 4: Graph of interconnecting IDs for multilingual equivalent of "city wall" and Object ID schema

From these three levels, the typology ontology model has been extended as an ID that joins the instance of the two complementary schemas (Figure 4).

2.3 Lexical Ontology Schema

The lexical knowledge model of building typology ontology is represented by sets of monolingual terms with lexical attributes implementing the Protégé knowledge representation tool in RDF. The Protégé tool is a knowledge-based editing object-oriented environment that supports the development of ontologies (http://protege.stanford.edu/). This editor supports the construction in a frame-like fashion with the classes and properties. The classes can be used to define the concepts in the ontology in a sub class-super class taxonomic manner and the slots or properties can contain values or information attached to the classes and their related instances (Noy et al., 2000). The lexical ontology designed in the Protégé consists of terminology schema and multilingual taxonomy schema.

2.4 Terminology Schema in RDF

A term set for each language was collected in the separated classes (Figure 5). Each term is the name of a building type in the historical citadel. The lexical attributes of the terms are defined as the slots or properties (or relationships) between the instances of these classes. Part of the RDF schema is shown in appendix of this paper.

The lexical relationships were separated into two main groups as Lexical Ontology Schema, Attributes Language-Dependant consisting of language dependant attributes as properties created separately for each class of the multilingual term-set. For example "*word*" with value type string; and Lexical Ontology Schema, Attributes Language-Independent consisting of language independent attributes as sets of properties for the relationships defined as lexical attributes.

Each one is created as a shared property for all the classes of term-sets of the different languages as follows:

"*Has morphology of*": with the value type **class** and allows super class **morphology**. This slot provides information related to the term associated to the morphological attributes, such as single words, complex words, or word for word translations.
"*Has Phonology of*": with the value type **string (**A property containing a string or a text), which contains the pronunciation of each term.
"*Has semantic of*": with value type **instance (**a property containing instances that are connected to other properties), which links all the attributes related to a semantic or the meaning of a term listed as:
"*Has description of*": with the value type **instance**, this links the definition of a term in the related bibliographic references;
"*Has synonym of*": with the value type **instance**, this takes synonym terms as the instances from the term-sets;
"*Has etymology of*": with the value type **instance**, this links a set of etymological attributes, such as old terms, depending languages, depending years, or specifications;
"*Has reference of*": with the value type **instance**, this links a set of properties with the value type **string** with the bibliographic information, such as the name, author, publisher, place, or ISBN;

The relationships in the lexical knowledge model described above were defined by the manual mapping from similar attributes in large dictionaries such as Oxford (Simpson, 1989) for the relationships as *definitions*, *etymologies*, *references*, etc. Lexical ontologies such as Word Net (Miller, 1995) and GOLD (GOLD, 2008) also support part of the lexical relationship schema, such as the *synonymy*, *phonology,* and *morphology*.

We have made a survey on the theoretical references related to the case study in particular those cover names of the buildings. This consisted of the publications related to the history of the region and the Citadel of Bam, the annual reports of the conservation project of the related organization (ICHHTO), and even historical travelogues (e.g. Ibne-Haugal, 937-977 A.D.). We have made an oral survey on the site to collect any related information from domain experts. According to the field study and the theoretical data collection, around 50 building types are recognized and inserted as instances of the class terminology monolingual terms.

3. MULTILINGUAL EQUIVALENT & THESAURUS

3.1 Multilingual schema

In order to provide the multilingual equivalents for a term, we first defined the separated classes of monolingual terms in the target languages of Persian, English, Japanese, and French. Then, we instantiated the schema described above with terms representing building types for each language.

We later designed a separate class for the multilingual equivalents. Each set of equivalents of the different instances of terms were linked with the slot "*has multilingual equivalent of*" with a value type instance. The allowed subclasses of this slot were the terminology, which means it takes instances as the multilingual equivalents of a term from the monolingual term-set in different languages, as shown in figure 4 for the term "city wall".

We used the multilingual term sets based on a bridge ID structure. Each ID as an instance is connected to monolingual terms (Figure 6). No language was used as a bridge between the other languages because a term can possibly have an inexact equivalent in other languages due to the meaning and usage restrictions. Therefore, our multilingual schema consists of 50 instances that link the equivalents of a term for the four languages together (Figure 7).

3.2 Constraints of Multilingual Schema

During the process of providing a multilingual term-set for the building types of the case study for this research we confronted the following challenges:

As there is no unified reference of dictionaries for collecting technical terms in the domain of historical architecture in the Middle East, references such as glossaries related to historical architecture in various languages, such as English (Ching, 1995) and Persian (Fallahfar, 1379), have been studied, but in a time consuming manual process. The practical field survey of the site of Bam enabled us to extract the proper terms for this research.

In some cases finding the proper multi-lingual equivalents for the technical architectural typology terms faced ambiguity according to usage. No equivalence could be found in other languages for a special term that names a special building in this historical citadel. We give two examples of such terms here with the translation of their definitions from Persian as follows:

Tekkiyeh: A building for a special kind of religious ceremony, consisting of a central courtyard and platform or rooms around it, which is made based on two iwans* or a four iwans organization in the plan.
Zoorkhaneh: A building for practicing traditional Persian sport with a high dome-shaped central space.

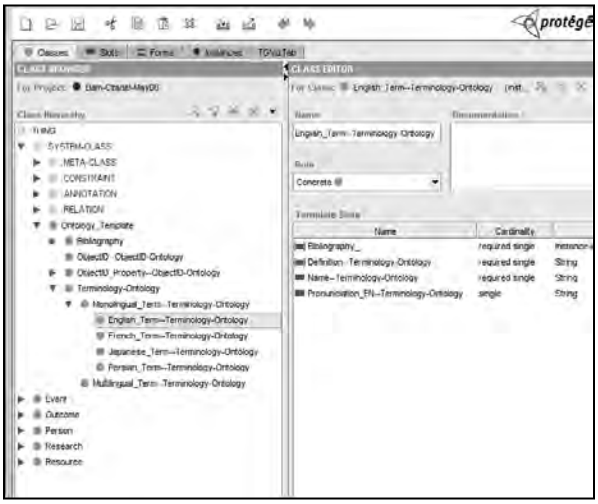


Figure 5: Hierarchy of classes and terminology schema in Protégé tool

Some of the standards for the multilingual equivalents or related references, such as the GOLD or ISO standard 5964 (ISO5964, 2002), can assist in the processing of the multilingual equivalents by defining the exact, inexact, partial, or single to multiple equivalents. One issue is the lack of multilingual equivalence standards for the field of historical architecture. The main target of this research for multilingual equivalence is to reach to a consensus for naming a building, in such cases the name is translated word for word and we have provided definition in the mother language and translation of it for others.

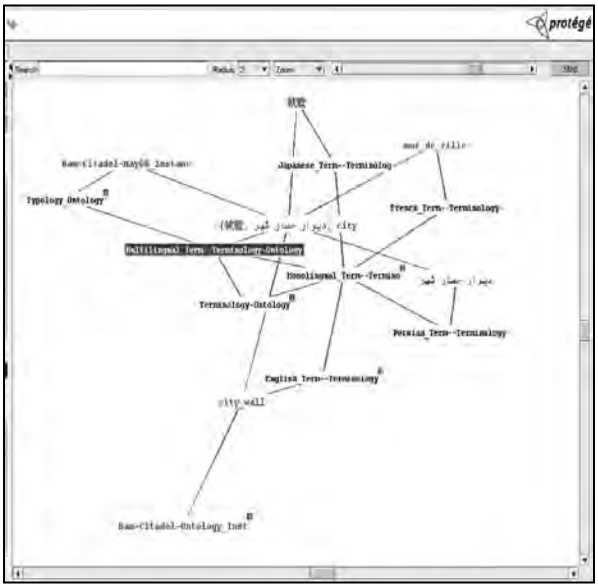


Figure 6: Graph of multilingual equivalents of term "city wall" that are connected by their related class by IDs in Protégé TGViz Tab

3.3 Taxonomy of Thesauruses

A terminology with an alphabetical order represents the knowledge not beyond that of a dictionary. Thesauruses give hierarchical structures to terminology to disambiguate the meaning of a term. Some general thesauruses, such as Wordreference (http://www.wordreference.com/) or visual thesaurus (http://www.wordreference.com/) are suitable for common terms, but for the architectural domain AAT seems to be a more suitable source for English terms. The hierarchy of AAT can be formalized as taxonomy of the classes and each term as an instance of terms. AAT online is not open source so the public access to hierarchy of classes of AAT as part of an ontology is not possible unless the licence agreement is obtained.

There is a lack of some technical domain terms with no equivalent in English in the AAT schema (such as Tekkiyeh). Besides thesauruses are language dependant; because the taxonomy proposed by the thesauruses is related to the meaning of the concepts or terms, and the hierarchical structure of thesauruses is cultural-dependant (ISO 5964, 2002). Although AAT is a powerful reference for the English language, Persian terminology cannot be properly matched with the English thesauruses. Furthermore, there is no Persian language thesaurus, nor could we find proper ones for Japanese language. For the French language, the database thesaurus of the Ministry for the Culture and Communication-Direction of Architecture and Inheritance can be proposed (http:// www.culture.gouv.fr/ public/mistral/thesarch_fr).

* A hall with barrel vault that is located in the middle of a side of a central courtyard with one side open; a characteristic of Persian architecture

4. ASSESSMENT AND APPLICATION

The ontology model for the Bam typology formalized by the Protégé tool provides an electronic knowledge base for the domain. It helps experts access and share information, and enables the semi-automatic validation of ontological information.

The process of evaluating the ontology consists of checking the consistency inside the ontology design tools. Running the ontology for the dependant tools, such as an image annotator, is another way ensuring the technical correctness of the ontology.

4.1 Technical Ontology Evaluation

The consistency and correctness of the ontology for the Bam typology is checked with the help of the RacerPro tool*, an OWL reasoner and inference server for the Semantic Web. For this purpose, we have exported the RDF file into OWL by using the Protégé tool for testing purposes. This test is run while the OWL interface of ontology is open. From the OWL menu *run ontology test* or *consistency checking* is selected. In the menu that is open at the bottom of the pages of these classes, the properties or individuals that contain errors are listed with the target error. The inverse slot of *hierarchical structure in English* language for our ontology is the slot value which is not properly defined and was corrected inside the lexical schema.

4.2 Tool-based Application

The usefulness of Bam typology ontology can be tested during the process of application by other tools or systems. As we described above the two complementary terminologies and the Object ID schemas are part of an upper level ontology consisting of the knowledge concerning the outcome, research, resources, people, and events of the supporting project for the 3DCG reconstruction of the Citadel of Bam. Due to limitation of licence part of the taxonomy is filtered for public access.

Figure 7: Instances of multilingual terminology for Bam typology inside the Protégé tool

* RACER stands for **R**enamed **A**Box and **C**oncept **E**xpression **R**easoner. RacerPro is the commercial name of the software., from http://agraph.franz.com/racer/

The output file of the ontology will contain the underlying information for the high-level support for web site design and development. The typical design process after the preparation of domain ontology consists of specifying the navigation structures and composing the user interfaces, defining the layouts and presentation styles, and expressing the customization requirements. This type of website does not contain static HTML pages. Instead, every page can be generated based on information contained in the ontology. **

5. CONCLUSION AND FUTURE WORK

We proposed a metadata-based multilingual ontology for supporting the semantic of the building typology of a historical city, part of the UNESCO world heritage site in danger, the Citadel of Bam, which was damaged in an earthquake in 2003. Our approach covered the development of a multilingual terminology ontology that is supported by the taxonomy of a thesaurus, and a complementary schema for a metadata-based specification of each building as a term in a lexical schema. An Object ID metadata schema is used for this purpose.

Future work includes instantiating the ontology schema and metadata collection of the buildings in the site. We also plan to reach to a consensus with domain experts of the Bam typology about the content of terminology ontology for reaching a proper assessment. Furthermore problems to be considered are how to add multimedia content like Quick Time Virtual Reality movies of the digitally reconstructed citadel onto the webpage, and how to connect it to the ontology schema. Although the tool supports some multimedia content, such as images, movies, sounds or interactive maps are still problems to be solved.

REFERENCES

Andaroodi, E., 2006. Architectural Spatial Ontology Model, on a Corpus of Silk Roads Caravanserais for Advanced Classification, Ph.D. dissertation, the Graduate University for Advanced Studies, National Institutes of Informatics, Tokyo

Berners Lee, T., James, H., and Ora, L., 2001. The semantic web, a new web content that is meaningful to computers will unleash a revolution of new possibilities, Scientific American.com, Feature Article

Brickley, D. and Guha, R. V., (Eds.), 2000. Resource description framework (rdf) schema specification, Note: World Wide Web Consortium, W3C Candidate Recommendation, CR-rdf- schema-20000327, March 27.

Ching, F. D., 1995. *A visual dictionary of architecture*, WILEY, USA

Database thesaurus of the Ministry for Culture and Communication - direction of Architecture and Inheritance-Mérimée, Note: reference in French, from: http://www.culture.gouv.fr/public/mistral/thesarch_fr (accessed 12 August 2008)

Dean, M., Connolly, D., Harmelen, F., Henzinger, M. R., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., and Stein L. A., 2002. Owl web ontology language 1.0 reference, Note: World Wide Consortium, W3C Working Draft WD-wl-ref-20020729

** We envision opening the site to public in 2009

Fallahfar, S., 1379. *Terminology of Persian architecture*, Tehran, Kaamyaab, Note: Book in Persian.

General Ontology for Linguistic Description (GOLD), 2008. from: http://www.linguistics-ontology.org/gold.html (accessed 12 August 2008)

Gruber, T. R., 1995. Toward principles for the design of ontologies used for knowledge sharing, *International Journal of Human and Computer Studies,* Vol. 43, No. 5-6, Academic Press, Inc., 907-928

ICOM CIDOC, domain ontology for cultural heritage information http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34424 (accessed 12 August 2008)

Icon Class http://www.iconclass.nl/ (accessed 12 August 2008)

ISO 5964, Guidelines for the establishment and development of multilingual l thesauri, 2002. International Organization for standardization, from: http://www.collectionscanada.ca/iso/tc46sc9/standard/5964e.htm#1 (accessed 12 August 2008)

Knowledge Capture, Victoria, British Columbia, Canada, 2001. pp. 194-201.

Mehriar, M., 2003. The history of citadel of Bam, report of archaeological studies, Bam recovery office, Iranian Cultural Heritage and Tourism Organization.

Miller, G., 1995. Word net: a lexical database for English, Communications of ACM, Vol. 38, No.11

Noy, N. F., Fergerson, R., and Musen, M., 2000. The knowledge model of Protégé 2000: Combining interoperability and flexibility, *EKAW '00: Proceedings of the 12th European Workshop on Knowledge Acquisition, Modelling and Management*, Springer-Verlag, London, UK, pp.17-32.

Ono, K., Andaroodi, E., Einifar A., Abe, N., Matini, M., Bouet, O., Chopin, F., Kawai, T., Kitamoto, A., Ito, A., Mokhtari, E., Einifar, s., Beheshti, M., and Adle, C., 2008. 3DCG Reconstitution and Virtual Reality of UNESCO World Heritage in Danger, the Citadel of Bam, *Journal of Progress in Informatics ,*National Institute of Informatics, No. 5. 99-137

Peterson, T., 1994. Introduction to the art and architecture thesaurus, Oxford University Press

Rumbaugh, J. R., Blaha, M. R., Lorensen, W., Eddy, F., and Premerlani, W., 1990. *Object-Oriented Modelling and Design*, Prentice Hall; 1st edition.

Simpson, J. A. (Ed.), 1989. *Oxford English Dictionary (20 volumes)*, Oxford English press.

Soo, W., Lee, C. Y., Yeh, J. J., and Chen, C. C., Using sharable ontology to retrieve historical images, 2002. International Conference on Digital Libraries, *Proceedings of the second ACMIEEE-CS joint conference on Digital libraries*, Portland, Oregon, USA, pp. 197-198

Smeulders A. W. M., Hardman L., Schreiber G., and Geusebroek J. M., 2002. An integrated multimedia approach to cultural heritage e-documents, *4th Intl. Workshop on Multimedia Information Retrieval, in conjunction with ACM Multimedia 2002*, ACM, New York, NY

Thornes, R. and Bold J., 1998. *Documenting the Cultural Heritage*, the J. Paul Getty Trust.

Thornes, R., Dorrell, P., and Lie, H., 1999. Introduction to Object ID, Guidelines for Making Records that Describe Art, Antiques, and Antiquities, the J. Paul Getty Trust.

Wielinga, B., Schreiber, A., Wielemaker, J., and Sandberg, J., 2001. From thesaurus to ontology, *International Conference Knowledge Capture*, Victoria, British Columbia, pp. 194-201.

Suggested Upper Merged Ontology http://sigmakee.cvs.sourceforge.net/*checkout*/sigmakee/KBs/Merge.kif (accessed 12 August 2008)

The Protégé editor and knowledge acquisition system http://protege.stanford.edu/ (accessed 12 August 2008)

United Nations standard products and services code (UNSPSC) http://www.unspsc.org/ (accessed 12 August 2008)

VRA 3.0 Core Categories http://www.vraweb.org/vracore3.htm(accessed 12 August 2008)

Visual Thesaurus http://www.visualthesaurus.com/?ad=google.dictionarysoftware.banner(accessed 12 August 2008)

Word Net http://wordnet.princeton.edu/perl/webwn(accessed 12 August 2008)

Wordreference http://www.wordreference.com/(accessed 12 August 2008)

ACKNOWLEDGEMENTS

The supporting research project, 3DCG reconstruction of the Bam and its cultural landscape is a collaborative project between Digital Silk Roads project of NII and ICHHTO. We would like to appreciate the scientific support of Associate Professor Dr. Frederic Andres for ontology development. This work was conducted using the Protégé resource.

APPENDIX A

RDF ontology file of part of terminology schema for the Citadel of Bam for Persian and English terms (designed by the Protégé) is presented as below:


```
<?xml version='1.0' encoding='UTF-8'?>
<!DOCTYPE rdf:RDF [
        <!ENTITY rdf 'http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
        <!ENTITY a 'http://protege.stanford.edu/system#'>
        <!ENTITY kb 'http://protege.stanford.edu/kb#'>
        <!ENTITY rdfs 'http://www.w3.org/2000/01/rdf-schema#'>
]>
<rdf:RDF xmlns:rdf="&rdf;"
        xmlns:a="&a;"
        xmlns:kb="&kb;"
        xmlns:rdfs="&rdfs;">
<rdfs:Class rdf:about="&kb;English_Term--Terminology-Ontology"
        rdfs:label="English_Term--Terminology-Ontology">
        <rdfs:subClassOf rdf:resource="&kb;Monolingual_Term--Terminology-Ontology"/>
</rdfs:Class>
<rdf:Property rdf:about="&kb;Explanation_of_Historical_References"
        rdfs:label="Explanation_of_Historical_References">
        <rdfs:domain rdf:resource="&kb;Persian_Term--Terminology-Ontology"/>
        <rdfs:range rdf:resource="&rdfs;Resource"/>
</rdf:Property>
<rdfs:Class rdf:about="&kb;Monolingual_Term--Terminology-Ontology"
        rdfs:label="Monolingual_Term--Terminology-Ontology">
        <rdfs:subClassOf rdf:resource="&kb;Terminology-Ontology"/>
</rdfs:Class>
<rdf:Property rdf:about="&kb;Name--Terminology-Ontology"
        a:maxCardinality="1"
        a:minCardinality="1"
        rdfs:label="Name--Terminology-Ontology">
        <rdfs:domain rdf:resource="&kb;English_Term--Terminology-Ontology"/>
        <rdfs:domain rdf:resource="&kb;Monolingual_Term--Terminology-Ontology"/>
        <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdfs:Class rdf:about="&kb;Ontology_Template"
        rdfs:label="Ontology_Template">
        <rdfs:subClassOf rdf:resource="&rdfs;Resource"/>
</rdfs:Class>
<rdfs:Class rdf:about="&kb;Persian_Term--Terminology-Ontology"
        rdfs:label="Persian_Term--Terminology-Ontology">
        <rdfs:subClassOf rdf:resource="&kb;Monolingual_Term--Terminology-Ontology"/>
</rdfs:Class>
<rdf:Property rdf:about="&kb;Pronunciation_EN--Terminology-Ontology"
        a:maxCardinality="1"
        a:minCardinality="1"
        rdfs:label="Pronunciation_EN--Terminology-Ontology">
        <rdfs:comment>Reference: http://www.merriam-webster.com/</rdfs:comment>
        <rdfs:domain rdf:resource="&kb;English_Term--Terminology-Ontology"/>
        <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdf:Property rdf:about="&kb;Pronunciation_FA--Terminology-Ontology"
        a:maxCardinality="1"
        a:minCardinality="1"
        rdfs:label="Pronunciation_FA--Terminology-Ontology">
        <rdfs:domain rdf:resource="&kb;Persian_Term--Terminology-Ontology"/>
        <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdf:Property rdf:about="&kb;Reference_of_Pronunciation_EN"
        a:maxCardinality="1"
        a:values="http://www.merriam-webster.com/"
        rdfs:label="Reference_of_Pronunciation_EN">
        <rdfs:domain rdf:resource="&kb;English_Term--Terminology-Ontology"/>
        <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdf:Property rdf:about="&kb;Reference_of_Pronunciation_FA"
        a:maxCardinality="1"
        a:values="http://en.wikipedia.org/wiki/Persian_phonology"
        rdfs:label="Reference_of_Pronunciation_FA">
        <rdfs:domain rdf:resource="&kb;Persian_Term--Terminology-Ontology"/>
        <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdfs:Class rdf:about="&kb;Terminology-Ontology"
        rdfs:label="Terminology-Ontology">
        <rdfs:subClassOf rdf:resource="&kb;Ontology_Template"/>
</rdfs:Class>
</rdf:RDF>
```

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

SUSTAINABLE MARKUP AND ANNOTATION OF 3D GEOMETRY

René Berndt[a], Sven Havemann[a], Volker Settgast[a], Dieter W. Fellner[a,b]

[a] Institute of Computer Graphics and Knowledge Visualization (CGV), TU Graz, Austria,
(r.berndt,s.havemann,v.settgast,d.fellner)@cgv.tugraz.at
[b] Fraunhofer IGD and GRIS, TU Darmstadt, Germany

KEY WORDS: annotation, markup, collada

ABSTRACT:

We propose a novel general method to enrich ordinary 3D models with semantic information. Based on the Collada format this approach fits perfectly into the XML world: It allows bi-directional linking, from a web resource to a (part of) a 3D model, and the reverse direction as well. We also describe our software framework prototype for 3D-annotation by non-3D-specialists, in our case cultural heritage professionals.

1. INTRODUCTION

The contribution of this paper is a simple solution to a seemingly simple, but in fact quite complex problem: To attach semantic information to 3D objects in a sustainable way. The complexity of the problem is a consequence of the generality of the required solution: There are numerous ways of representing 3D-objects, and there are numerous ways to attach information. The sustainability requirement implies the use of well-established standards, and simplicity is a must in order to make it very easy for other applications to support and implement our proposed attachment technology. We desire, and anticipate, wide-spread use of the described solution.

1.1 Motivation: The Arrigo Showcase

The original motivation of our work came from a very well-made presentation of cultural 3D-objects, the *Arrigo Showcase* from our colleagues around Paolo Cignoni and Roberto Scopigno from ISTI-CNR in Pisa (Callieri et al., 2006),(Callieri et al., 2008). It represents a prototype case of cultural information visualisation to a public audience today. It uses the VirtualInspector library for interactive display of massive multi-resolution meshes. The software is shown in a 3D-kiosk located next to the real statues, so that museum visitors can at the same time appreciate the beauty of the real, and enjoy additional information attached to the virtual statues (see Fig. 1).

The truth, however, is that the Arrigo Showcase, how beautifully it may be made, is a **dead end** in terms of knowledge management: To really look good, the information pane in Fig. 1 (right) is a bitmap created by a web designer. The point is, it is *only* stored as a bitmap on the Arrigo DVD. Fortunately, the original text was eventually found by our colleagues; but even the text is not very helpful in terms of sustainable knowledge management.

Sustainability can be achieved only by converting the facts in the text to some form of semantic format, amenable to automatic processing. One such format, developed for the cultural sector, is CIDOC-CRM, the *Conceptual Reference Model* of the *International Committee for Documentation of the International Council of Museums* (ICOM-CIDOC) (Crofts et al., 2005). The CRM (version 4.2) defines 84 *classes* for entities such as actor (person), place, event, time-span, man-made object, etc., and 141 relations between those entities, such as participated in, performed, at some time within, took place at, is referred to by, and has created. Using the *Resource Description Framework*, CIDOC-CRM facts can be encoded as RDF-triplets (subject-predicate-object, formally entity-property-entity), in which case each entity has a type and refers unambiguously to an object. This unambiguous reference is simply encoded as *Uniform Resource Identifier* (URI), basically a URL. It is supposed to be a persistent identifier, a unique ID, that may, but does not have to, point to an existing web resource. In that sense the URI represents a sustainable piece of knowledge. This reduces the problem of attaching information to the problem of attaching a URI to a 3D model.

Figure 1: Arrigo Showcase: Museum visitors can interactively navigate (rotate, pan, zoom) a 3D model and obtain additional information by clicking on information links embedded with the 3D model.

1.2 Attaching URLs to 3D-Objects

Attaching a URL to a (part of) a 3D-file in general is a difficult problem, as was noted already by a number of researchers, just to mention Jarek Rossignac (Rossignac, 1997). One of the most annoying problems is the 3D file format problem: Hundreds of competing 3D-formats exist, each with its pros and cons, and with its own community whose members usually do not care about other communities. This leads to fragmentation and incompatibility: Painful expierence teaches that almost any 3D format conversion inevitably implies **loss of information**. This leads to much friction and avoidable cost especially in the professional *digital content creation* (DCC) industry. To facilitate content production for the Playstation 3, Sony Entertainment consequently initiated the Collada initiative which came up with a new XML-based exchange 3D-format, which is now an open format hosted by the Khronos group (Arnaud and Barnes, 2006).

Since our concern is knowledge exchange, we chose Collada as the main format. It has a number of quite desirable properties for our purposes:

- **Clear, simple structure**: A Collada file has a ⟨library⟩ part, where the content nodes (models, materials etc) are defined, and a ⟨scene⟩ part with the scene hierarchy and transformations, which typically references library nodes by relative (intra-file) links
- **Relative anchors and links**: Library nodes can automatically be referenced using relative links and, thus, there is a way of adressing them from external resources using the #-syntax
- **No embedded BLOBs**: Collada files may not contain binary content, thus, there is a clear separation between structure and content; no massive datasets within Collada
- **Reference to external files**: Massive datasets can be referenced as external geometry; the file name is in fact a URL, which permits to include remote datasets transparently
- **Sustainable annotations**: Almost any type of node can have an ⟨extra⟩ child node attached that may contain arbitrary information encoded in XML. An application may claim itself Collada compliant only if saving a file leaves the ⟨extra⟩ information of a loaded object untouched.
- **Wide-spread use**: Collada is the 3D format of Google Earth and Google Warehouse, which offers masses of 3D content for free, and with Google Sketchup an authoring application for non-expert users is also freely available.

As pure ASCII format Collada is useful for encoding the low-poly models for games, or the clean CAD-like Sketchup models, but certainly not for massive scanned datasets with billions of triangles. Compressed triangle mesh formats typically need only 2-5 bytes per vertex (topology *and* geometry), much less than, e.g., compressed Collada ASCII. Furthermore, in Cultural Heritage certain custom shape representations are used which integrate, e.g., information for image-based rendering (BTF and PTM), for which special file formats exist. So the solution is to use Collada merely as a scene structuring format (we call it "*Collada light*") with references to external *geometry files* containing 3D-data in one of several well-documented supported formats.

The great advantage is that since the annotations are stored in the Collada file, the original 3D datasets remain untouched. Furthermore, the same objects when stored, e.g., on a web server, can appear in different Collada files; and possibly also with different annotations. This is consistent with the view that semantic information is, in some way, always an interpretation. It is even possible that researchers exchange such lightweight Collada files via e-mail, for instance to discuss different interpretations of an artefact. When the file is opened, the scanned dataset is retrieved from a central server, if requested even in a reduced resolution.

1.3 An easy-to-use practical solution

Now the question is: Which types of annotations are possible, and how can researchers or curators create these annotations easily?

Shape markup does not always have to be carried out manually. One could also employ automatic feature detection algorithms that infer, by statistical evidence, certain types of semantic information. Algorithms exist for wall detection, for detection of symmetries, for detecting part-of relations (instantiation) and similarities between shapes, for primitive decomposition (sphere, box, cylinder etc.), as well as for surface properties such as curvature, roughness, and the like. All these algorithms were published in scientific papers, but no method exists to reliably and sustainably store, combine, and explore their results. Our proposition is to create easy-to-use interfaces for these algorithms, and to store their results as semantic attachments in Collada.

This peek into the future already indicates that a flexible type of software infrastructure is needed, that is very versatile and proves useful in a whole range of different applications. We propose to use a component architecture, in our case Microsoft's ActiveX, to create a very simple widget that merely consists of a rectangular OpenGL drawing canvas. This ActiveX control, called *ActiveEpochViewer*, may be embedded into applications that can be created using easy-to-learn languages like C# or VisualBasic that are useful for GUI design.

The ActiveX control contains an open source scene graph engine, in our case OpenSG, that fosters sustainability by being well supported and extensible: When a new 3D format is to be included, it is sufficient to create a loader and new type of OpenSG geometry node for rendering. On the Collada level, the 3D-format is completely transparent, since it is just a URL.

The last missing ingredients are the communication between ActiveX control and application, and a method for rich 3D interaction (behaviour and events). In our case, both are handled by the *Generative Modeling Language* (GML) (Havemann, 2005): When a GUI button is pressed, VisualBasic (for example) sends a pre-defined string with GML commands to the ActiveEpochViewer, to reposition the camera, to switch to attachment mode, or to load a file etc. It can, however, also attach a GML-callback to a 3D object. This callback is to be executed, e.g., when the object is clicked. Since GML is a full scripting language that has access to the whole scene graph, arbitrarily complex actions can be issued this way, leading to rich 3D interaction.

We have realized two example applications based on this method, one is an **authoring application** for creating and editing shape markup, the other is a **presentation environment** to explore the resulting semantically enriched shapes.

2. RELATED WORK

To the best of our knowledge, our work is novel in that no previous work exists which focused on solving the general problem of sustainable markup and annotation of 3D geometry. However, we are of course not the first ones to attach information to 3D objects. Numerous projects exist where, like in the Arrigo Showcase, custom and ad-hoc solutions for 3D-annotation were invented. And numerous proprietary solutions also exist in the context of *product lifecycle management* (PLM) in the automotive, aerospace, and building industries, where 3D objects can also be versioned on a sub-part level, and every change is documented pedantically for a long time with author and clearing date.

But also in these domains, the awareness increases that a sustainable knowledge management is still missing. With liability periods of 30 years and more, manufacturers need reliable standards for vendor-independent markup of 3D geometry, to become robust against software version changes. On the European level for instance, the MACE project with its upcoming conference look into sustainable metadata standards for architecture. Our focus, however, is to devise a mechanism how to attach information to 3D objects in general, independent of a particular domain.

One inspiration for our work is the *Digital Photo Library* of the *Kunsthistorisches Institut in Florenz*, freely accessible under (khi.fotothek.org), where researchers can browse historic images. When they mark a spot on an image, the result is a URL that they can send to a fellow researcher, for instance to make him look at a statue on the facade of the dome in Florence:

http://digilib2.gwdg.de/...&mk=0.1179/0.5021&pt=0

Links can be attached to 3D-objects in the VRML/X3D file formats since a long time. However, links are attached only on a per-object basis, and there is no standard way of attaching markup geometry as an annotation to any node. And we found Collada easier to use for our purposes than X3D, especially in the "light" version with external content formats. – One attempt to attach semantic information to X3D files was made by Franco Niccolucci and Andrea D'Andrea (Niccolucci and D'Andrea, 2006). However, the linking was not bi-directional, and focus was the object ontology.

We would like to note that there is, of course, a standard XML technology for bi-directional linking between 3D objects and semantic information, namely XPointer / XPath / XLink. However, simplicity is an asset, and since our Collada light-files have only the duty of storing the scene hierarchy and semantic annotations, we did not see the point in treating each link as a separate entity of its own. Furthermore, to address objects on a sub-object level, we simply need a geometric target for a link in the file, which is exactly what our markup geometry is.

Our work can maybe best be understood in the context of an article about seven open research problems with 3D objects today (Havemann and Fellner, 2007):

- Classify all shape representations in CG
- A sustainable 3D file format
- Generic, stable, and detailed 3D markup
- Generic 3D query operations
- Paradata: Processing history, provenance
- Close the semantic gap: Meaning of shape
- Maintain consistent relation between shape and meaning

The present paper can be understood as a first solution to the third problem.

3. INFORMATION LINKING

This section will describe the first technical ingredient of our system, the annotation format used with the Collada ⟨extra⟩ tag, that is then evaluated in OpenSG. OpenSG is a powerful and extensible realtime graphics system and scene graph engine, capable of displaying many kinds of 3D object representations. It also offers the possibility of attaching arbitrary data to scene graph elements. We use this technique to store additional information directly within the scene. The annotations are attached as text strings that contain unmodified XML code.

The additional information is represented as a combination of URLs, a markup geometry and a camera definition. The concept of URLs allows to stick any data to an annotation as long as a URL exists to find it. A single markup can contain any number of URLs. For the first version of this annotation concept a simple sphere shape is used as markup geometry. Other shapes are possible, but the simple sphere definition is one of the fundamental ways to mark a spot in 3D. More complex annotation geometry will be introduced after we have made some experience with spherical annotations.

Note that annotations are very often view dependent. Automatic ways of setting the camera transformation exist, but in general it is more reliable to let a human set up the point of view. Therefore at least one camera transformation is added to the data set, per default the camera set up at the time of annotation.

3.1 Format Specification

The following format specification for ⟨annotation⟩ can be part of any ⟨extra⟩ element, which belongs to a ⟨node⟩ within a Collada file. Each ⟨node⟩ element can contain multiple ⟨annotation⟩ tags in the ⟨extra⟩ part.

annotation

The ⟨annotation⟩ element represents one annotation attached to the ⟨extra⟩ element of a node within the scene. The annotation defines a geometry, one or more camera definitions, and one or more information links.

Attributes

id	xs:ID	A text string value containing the subidentifier of this element. This value must be unique within the scope of the parent element. Required.
title	xs:string	A text string containing the title of the annotation. Optional.

Child Elements

Name	Description (Occurrences)
⟨annotation_geometry⟩	Geometry defining the bounding of the annotation. (1)
⟨annotation_camera⟩	Camera position (0 or more)
⟨annotation_url⟩	Unique identifier to the additional data (1 or more)

annotation_geometry

The ⟨annotation_geometry⟩ defines the position of the annotation with in the scene. At the current implementation level the geometry can consist of one or more spheres, but arbitrary other geometry (like oriented boxes or free-form shapes) would be possible in the same way.

Child Elements

Name	Description (Occurrences)
⟨annotation_sphere⟩	Geometry defining the bounding of the annotation. (1 or more)

annotation_camera

This element defines a camera for the parent ⟨annotaton⟩. It is used to move the camera to a reasonable position when viewing the annotation.

Child Elements

Name	Description (Occurrences)
⟨from⟩	A point in space to look from. (1)
⟨at⟩	The camera target position (1)
⟨up⟩	The up-vector of the camera. (1)

annotation_url

Unique identifier (URL) to the additional data.

title	xs:string	A text string containing the title of the information source. Optional.

annotation_sphere

Describes a bounding sphere, which is used to define the geometric reference of the annotation within the current node.

Child Elements

Name	Description (Occurrences)
⟨center⟩	The center of the sphere. (1)
⟨radius⟩	The radius of the sphere (1)

An example of a Collada attachment where these nodes are used is shown in Fig. 2.

```
 ...
<annotation  id="AnchorId"
             title="Additional_Information">
  <annotation_geometry>
    <annotation_sphere>
      <center>12.9  33.94  3.12</center>
      <radius>3.84</radius>
    </annotation_sphere>
  </annotation_geometry>
  <annotation_camera>
    <from>44.8  34.8  28.58</from>
    <at>7.5  35.6  −1.64</at>
    <up>0.02  0.99  0.05</up>
  </annotation_camera>
  <annotation_url  title="Collada">
   http://www.khronos.org/collada/
  </annotation_url>
  <annotation_url  title="OpenSG">
   http://opensg.vrsource.org/trac
  </annotation_url>
</annotation
 ...
```

Figure 2: Example ⟨annotation⟩: The annotation "AnchorId" with one sphere as the annotations markup geometry, a camera definition and links to two external information sources.

4. EPOCH-VIEWER FRAMEWORK

The core of the Epoch-Viewer framework is a reuseable component, which combines OpenSG with the Generative Modeling Language (GML) (Gerth et al., 2005). The GML was introduced by Havemann (Havemann, 2005) and provides a stackbased language (similar to Adobe's Postscript). Most of the functionality of the OpenSG API was exposed to the GML language. Since the first version the list of available operators has been continuously extended, e.g., support for reading the Collada scene description and scripted XML code manipulation was added.

In the current version this component is realised as an ActiveX control, which is compatible with the interface of the ActiveGML plugins (Berndt et al., 2005). The following C# snippet shows the typical usage of the control:

```
...
EpochViewer.LoadModel("mark1.xgml");
EpochViewer.Call("mark1.main");
...
```

One feature of the new ActiveEpochViewer control is its concise API with only four methods, two of which were available in the original ActiveGML methods (marked *italic*):

- *LoadLibrary(string)*
- *Call(string)*
- GetStackSize()
- GetStackElementAsString(index)

The two new methods replace the depreciated getStackTopAsString. The limitation of the existing interface was that only communication from the container application to the ActiveX control was possible. With the new introduced event called GMLEvent a bi-directional communication can be established. So the container application can receive notification from the scene graph engine (e.g. a node has been clicked). A complete technical overview of the Epoch-Viewer framework would go beyond the scope of this proposal and will be described in a separate article.

This component can be used by all ActiveX-compliant software (like Microsoft Office, Internet Explorer, .NET Framework), which makes it the ideal plattform for generating tailor-made graphical applications. Figure 3 shows the control embedded in a web page.

Figure 3: The Epoch-Viewer ActiveX control inserted into an HTML page using the object tag. GML functions for loading and highlighting the annotations are triggered with Java Script.

5. SAMPLE APPLICATION

The Epoch-Viewer Framework is used by two showcase applications, an exemplary authoring tool and a viewer for the created semantically enriched 3D objects. The annotation procedure should be as easy as possible. People who have special knowlegde of an object, e.g. a historian or an archaeologist, will not always have a lot of experience in 3D applications. So we tried to support non-3D-expert users by minimizing the functionality to what is absolutely necessary.

For the authoring application a standard mouse navigation is used. The user orbits around the scene and is able to zoom and pan to find good observation points for the annotations. In presentation mode, other forms of navigation are also supported including joystick input and tracking devices. Navigation is fully customizable by scripting.

A GML script defines internal functionality and user interaction. The GUI elements are capsulated into the enclosing .NET application. Smaller changes to the functionality can be archieved by modifying the GML script alone. For the showcase applications there are four functions in the script that are triggered by the container form.

- `load-file` and `save-file`
- `show-anchor`

Figure 4: The showcase for authoring: The user navigates to a position of interest. Annotations are added by clicking on the model and dragging the radius of the markup sphere. In the anchor form multiple URLs can be attached by typing or dragging the adress from a browser. The dialog also includes a browser for preview and web navigation. (web browser: www.wikipedia.org)

- `update-data`

For quick processing the anchors of a scene are stored in a list. In GML, anchors are represented as dirctionaries.

`load-file`

The `load-file` function expects a filename or URL of the selected scene on the interpreter stack. A standard open file dialog can be used in the container form to get a valid filename. The filename is pushed on the stack and `load-file` is called.

Asynchronous loading is done by the loading thread in the OpenSG framework. A callback function is defined and connected to the load finish event. After the loading process is done, this call back function is triggered. A scene may consist of a single node or a tree of nodes.

With the new node as entry point, the subscene is traversed. If there are geometry nodes containing XML attachments, they may already contain annotations. The XML code is analysed in another GML function. If there is an ⟨extra⟩ element, `parse-extra` is called to extract annotation information.

For the definition of new annotations every geometry node gets a mouse call back function. If there is a click on the geometry, a new markup sphere is generated. By dragging the mouse, the radius of the sphere is defined. The mouse button release event triggers an event back to the editor application for the annotation enter dialog. After the user has finished editing the annotation, the editor calls `update-data`.

The container form has a save as button that opens a standard save dialog and a save button. For the save as function, the form pushes the filename onto the stack and calls `save-file-as`, similar to the load function. The `save-file` function is called for the save button and triggers a write procedure in the control to save the scene back into the same Collada file.

`show-anchor`

The `show-anchor` function is called from the container form by the time the user requests the annotation data. This is possible by clicking on one of the anchors in the anchor list. The function expects the name of the anchor on the stack. If the anchor is found in the list of anchors, the markup geometry is highlighted and the camera animation is started to move to the defined location.

`update-data`

This function is called by the container form every time the user has modified the annotation in the anchor dialog. It expects the new XML code and the old anchor name on the stack. The old data is deleted from the list of annotation dictionaries and the new data is added. The old anchor name can be replaced by a new one defined in the dialog. The last step is to attach the new XML data to the OpenSG node.

`parse-extra`

The `parse-extra` function seeks the XML code of the ⟨extra⟩ element for ⟨annotation⟩ elements. For every annotation element, the annotation geometry is created. In our example code, this is always `add-sphere`.

`add-sphere`

This function gets the position and the radius of a new markup sphere and calls the primitve creation provided by the OpenSG control. The geometry is added to the scene graph in a special node that is excluded from exporting. For the markup geometry, again call back functions have to be defined to react on user mouse clicks. If the sphere is clicked, a camera animation moves the point of view to the predefined location and a special FinishPoint event is fired. The container application reacts on this event by adding a new anchor element to the internal anchor list.

The Viewer

For the viewer application the same GML code can be used for highlighting annotations and animating the camera. Depending on the purpose of the presentation GML scripting is a flexible way to generate all kinds of showcases as introduced in (Havemann et al., 2007). Besides the Epoch-Viewer control the container window includes a browser control to display the URL targets and a list of anchors (see Figure 5). Annotations can either be examined by clicking on the markup geometry in the 3D scene or by selecting an anchor from the list. By choosing an annotation, the point of view is automatically moved to the predefined position. Manual navigation within the 3D space is also possible.

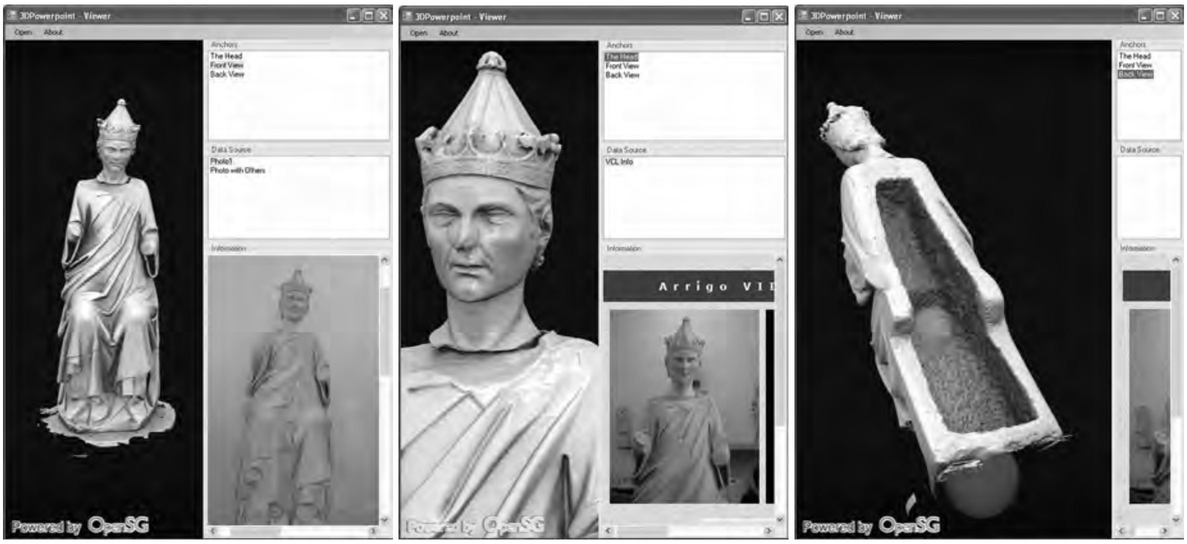


Figure 5: A simple viewer showcase: The user can examine annotations by clicking on the markup sphere in the 3D scene or by selecting the anchor name in the list at the top right of the application window. The camera automatically moves to the predefined observing position. A list of attached URLs is displayed for the currently selected anchor. A simple browser control can display the content of the URL's target. (images in web browser: www.scultura-italiana.com (left), vcg.isti.cnr.it)

6. CONCLUSION

As promised we have presented a simple solution to a complex problem: 3D objects, even massive scanned datasets, can be provided with semantic annotations to provide arbitrary types of additional information on a sub-object level. The head of a statue, even its nose, can be linked bi-directionally to web resources like web pages, pdf documents, images and videos, and even Flash presentations – anything a standard internet browser can display.

The format specification is also simple enough that automatic feature detection algorithms could be easily designed so that they produce Collada light files to let users browse through the detected features. The format is also suitable for long-time archival, for sustainable knowledge management, as long as it is made sure that the targets of the annotation links still exist; e.g., by using links that point into the same database. Consequently, one of our future research directions is to set up an extensive database of annotated Collada files, that point to web resources containing primary data (scanned datasets) of our colleagues in the EPOCH network of excellence on processing Cultural Heritage.

Another road for future research is to create more easy-to-use applications for creating, managing, browsing, and presenting semantic relations in 3D scenes. This has a huge potential since in our three-dimensional reality we are surrounded by man-made objects with a long-standing tradition, which are embedded in a dense network of semantic relations. Consequently, another research direction is to make these semantics explicit, by coupling the 3D exploration tightly with a semantic database. RDF-triplets encoding CIDOD-CRM knowledge could then be embedded live in the 3D scene.

Speaking of larger scenes, another research direction is the question of scalability. We envisage that when providing our authoring toolkit with the functionality for interactive scene assembly, things will soon get very complicated. Collaborative 3D markup is already technically possible now (even in a time-synchronized manner), but we are aware that consistency issues and resolving conflicts will then become a serious issue. With denser semantics, also creating different views on demand by database-like selection queries for particular semantics will be needed.

So, much remains to be done, in fact we have only opened the door to a hopefully fruitful and important new research field. Nevertheless we believe that the presented technologies are a sound foundation for future applications to stand on.

The software, both the ActiveEpochViewer control and the sample applications, can be downloaded for non-commercial use from *www.cgv.tugraz.at/EpochViewer*.

Acknowledgements

The authors want to thank Roberto Scopigno, Paulo Cignoni and the ISTI-CNR for the nice Arrigo models. And the authors wish to thank the European Commission, in particular the EPOCH network of excellence in cultural heritage (www.epoch.eu) for their generous support.

REFERENCES

Arnaud, R. and Barnes, M. C. 2006. *Collada: Sailing the Gulf of 3d Digital Content Creation*. AK Peters Ltd.

Berndt, R., Fellner, D. W., and Havemann, S. 2005. Generative 3d models: A key to more information within less bandwidth at higher quality. In *Web3D '05: Proceedings of the tenth international conference on 3D Web technology*, pages 111–121, New York, NY, USA. ACM Press.

Callieri, M., Ponchio, F., Cignoni, P., and Scopigno, R. 2006. Easy access to huge 3d models of works of art. In Fellner, D., editor, *Fourth Eurographics Italian Chapter 2006*, pages 29–36, Graz, Austria. Graz University of Technology, Eurographics Association. Catania, 22-24 February 2006.

Callieri, M., Ponchio, F., Cignoni, P., and Scopigno, R. 2008. Virtual inspector: a flexible visualizer for dense 3d scanned models. *IEEE Computer Graphics and Applications*, 28(1), pp. 44–54.

Crofts, N., Doerr, M., Gill, T., Stead, S., and Stiff, M. 2005. *Definition of the CIDOC Conceptual Reference Model*. CIDOC Documentation Standards Working Group, version 4.2 edition. Also ISO/PRF 21127, available from cidoc.ics.forth.gr.

Gerth, B., Berndt, R., Havemann, S., and Fellner, D. W. 2005. 3d modeling for non-expert users with the castle construction kit v0.5. In Mudge, M., Ryan, N., and Scopigno, R., editors, *Proc. VAST 2005*, 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, pages 49–57. Eurographics / ACM Siggraph, Eurographics Press.

Havemann, S. 2005. *Generative Mesh Modeling*. PhD thesis, Institute of Computer Graphics, Faculty of Computer Science, Technical University Braunschweig, Germany.

Havemann, S. and Fellner, D. W. 2007. Seven research challenges of generalized 3d documents. *IEEE Computer Graphics and Applications*, 27(3), pp. 70–76. (special issue on 3D documents).

Havemann, S., Settgast, V., Krottmaier, H., and Fellner, D. 2006. On the integration of 3d models into digital cultural heritage libraries. In Ioannides, Arnold, Niccolucci, and Mania, editors, *VAST 2006 Project Papers*, 7th Inter. Symp. on Virtual Reality, Archaeology and Cultural Heritage, pages 161–169, Nicosia, Cyprus. Epoch publication.

Havemann, S., Settgast, V., Lancelle, M., and Fellner, D. 2007. 3d-powerpoint - towards a design tool for digital exhibitions of cultural artifacts. In Arnold, Niccolucci, and Chalmers, editors, *VAST 2007*, 8th Inter. Symp. on Virtual Reality, Archaeology and Cultural Heritage, pages 39–46, Brighton, UK. Eurographics Association.

Ioannides, M., Arnold, D., Niccolucci, F., and Mania, K., editors 2006. *International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, Nicosia, Cyprus. Eurographics Association.

Microsoft Corporation 2002. *Microsoft; Visual C# .NET Language Reference*. Microsoft Press, Redmond, WA, USA.

Niccolucci, F. and D'Andrea, A. 2006. An Ontology for 3D Cultural Objects. In (Ioannides et al., 2006), pages 203–210.

Rogerson, D. 1997. *Inside COM*. Microsoft Press, Redmond, WA, USA.

Rossignac, J. 1997. The 3d revolution: Cad access for all! In *SMA '97: Proceedings of the 1997 International Conference on Shape Modeling and Applications (SMA '97)*, page 64, Washington, DC, USA. IEEE Computer Society.

Cultural Heritage Resource Information Systems

PROFILING ARTEFACT CHANGES: A METHODOLOGICAL PROPOSAL FOR THE CLASSIFICATION AND VISUALISATION OF ARCHITECTURAL TRANSFORMATIONS.

I.Dudek [a], J.Y Blaise [a]

[a] UMR CNRS/MCC 694 MAP, 184 av de Luminy, 13288 Marseille, France - (idu, jyb)@map.archi.fr

KEY WORDS: Architectural heritage life cycles, knowledge modelling, information visualisation, dynamic SVG

ABSTRACT:

When studying heritage artefacts and trying to represent what we know of them, it is important to portray not only key moments in evolution of artefacts, but also processes of transformation. In this contribution, we introduce a methodological framework of description of artefacts' transformation and investigate the usability and efficiency of diagrammatic representation as a mean to visualize the above mentioned framework. A focus of our interest is the way artefacts get transformed. The methodological proposal presented identifies as a central notion: a *life cycle* - a sum of states and transitions following in succession- in what can be seen as a diachronic approach. We then introduce the diagrams proposed in order to visualise the above mentioned life cycles and provide examples on major or minor architecture within the medieval part of the town of Kraków (former capital of Poland, experimental set for this research). Two types of diagrams are introduced: *diachrograms* that distribute along a time axis transitions and states, and *variograms* that detail the nature of the changes. A combination of these graphics should help better understanding, in a cleat-cut manner, how changes over time affect architecture. But it should also underline key aspects of data in "historical sciences": uncertainties, incompleteness, long ranges of time, unevenly distributed physical and temporal stratigraphy.

1. INTRODUCTION

1.1 General objective

Heritage artefacts, may they be individual edifices or whole sites, are rarely left unchanged by time. Various natural or man-related events occur throughout the centuries, resulting in numerous transformations. Such transformations can introduce minor architectural changes (extension, refurbishment, change of usage) or cause important modifications, both in terms of physical appearance and in terms of usage. Ultimately, what today we observe (the artefact itself) and know (historical analysis) can be understood as a collection of traces, traces of all the moments an artefact has been through during its often complex evolution. Methods exist that help researchers to state (although with remaining doubts) how an artefact has been at time t_1, t_2, …, t_n. Broadly speaking, such methods rely on "expert-interpretation" of observations or archival data. In addition, when possible, they can include cross-examinations of individual cases considered as similar on one or several aspect. But if one wants to represent and explain the processes that lead from state at time t_1 to states at time t_2, …, t_n, less solutions exist. As an answer, we investigate diagrammatic visual displays, that could foster a comparison-enabled, global vision of an artefact's evolution; whereas traditional architectural representation (CAD based or not, see (Estevez, 2001), by privileging shape modelling, tends to enhance states over changes. In other words, we here experiment the infovis path – amplify cognition (Kienreich, 2006), rather than the scientific visualization path – real object/graphics mapping (Spence, 2001).

Given an efficient, workable, identification of states and transitions, synthetic diagrams could offer an unprecedented view on the global evolution of an artefact. In addition, such graphics could possibly uncover patterns of evolution within a site or across sites, underline uncertainties or exceptions ("documentary gaps"), raise questions about the relative evolution of families of artefacts (urban houses in this or that quarter of the city, churches across the city, gothic castle across a wider territory, etc.). Accordingly, our objective is at the intersection of two issues:

- a methodological framework enabling the description of architectural changes (*i.e.* a knowledge modelling issue),
- a set of visual signs and/or diagrams developed in order to apply (and evaluate) the above mentioned framework on real cases (*i.e.* an infovis issue).

In this paper we will focus on the former aspect, with the latter used as a mean to evaluate it on real cases.

1.2 Context and background

When studying heritage architecture, the heterogeneous nature of documents handled poses a number of problems (*i.e.* ambiguity of textual descriptions, exactitude of artistic representations benefiting largely from *licencia artistica*, etc.). A resulting interpretation is drawn from the reading of a historical source, as well as from the experience, knowledge and intuition of the analysts. The result of this process is a hypothesis. In "historical sciences" however, one cannot use experiment in order to verify a hypothesis (an experiment in the past is not very realistic!). Therefore the only way to amplify a validity of a hypothesis is to confront it with unknown data or with the results of other scientists (*cf. intersubjectivity* (Bocheński, 1968)). However such cross-examinations should go beyond a parallel reading of states, and include an analysis and visualisation of causes and effects. Now, when analysing major current research axes, one can observe that *synchronic* approaches are strongly dominant.

Renewed survey techniques, simulation through 3D modelling and virtual reality, site management systems using GIS platforms, archival information systems are among the most prominent results, with at the end of the day a number of clues related to one or another moment in the evolution of an artefact.

On the “computer graphics and/or new technologies” side, various spatial granularities are observed (ranging from cities (Lerma, 2004) to architectural interiors (Perkins, 2003)). However, in the above mentioned field, architecture has served mainly as a test bench -here for immersive platforms, there for 3D surveying, etc.. A good example can be found in (Suverg, 2003) where urban architecture, illustrating a research on geometric reconstruction, is described only through three parametric building models called primitives: flat roof, symmetrical gable roof building, nonsymmetrical gable roof building (sic!). Another typical example is the use of VR for site presentation (Ando, 2003) where architecture finally appears as a context-free phenomenon, with a still-image of the edifice giving no hint on why, but unduly portraying a moment in the past (for which we have no such certainty as the image suggests). In the above examples, architecture as a domain is widely ignored, and the scientific conclusions would have been the same if scientists had studied ship containers, kennels, rabbit warrens, etc..

On the “humanities” side, a strong investment has been done in the past decade on computer-aided data recording, analysis and management (see (Müller, 1997), (Ramondino, 2001), (De Luca, 2005)), with common applications ranging for instance from expert-oriented archaeological site management (Sebillo, 2003), (Huber, 2000) to didactic end-user visual disposals (Kodym, 1999), (Kantner, 2000), (MhMK, 2007). But here again, the effect of time and events (that of course researchers are very aware of) is far from being stressed.

As can be seen from these examples, the focus is ultimately put on describing specific (and chosen) moments in the life of an edifice or a site, may this moment be contemporary (surveys, site management), or may it be time t_n of the artefact’s evolution (3D reconstruction, documentation and archive studies).Very little has specifically been done, in the field of the architectural heritage, in order to describe and represent visually the time-chain between successive states or moments in the evolution of artefacts. However such approaches can be found outside of the field of the architectural heritage, and in particular in geography. Time-geography (Lenntorp, 2003) is typically an approach where the focus is put on time motion. Still, time-geography applies to spatial concepts that do not match the granularity one faces in architecture. Moreover, time geography applies mainly to short life cycles, where the global spatial context (inside which the variations over time of a phenomenon are studied) remains unchanged.

A much closer example can be found in graphs proposed in (Renolen, 1997), where changes in land areas are visually assessed through synthetic, easy-to-read diagrams. Renolen describes and represents territorial changes: he isolates states and defines events causing changes, a view that we base on. However, his field of application is land areas as seen by a geographer, and the graphs proposed are not directly from being applicable to architectural changes. Among noticeable differences are the 3D nature of artefacts, the transformation processes within a given and stable land area, the reuses and displacements, long-term internments of built structures, uncertainties and incompleteness in the data sets, in the dating of events, in the actual physical transformations, etc.

In this paper, we develop an analysis of states and transitions that we believe better matches the specificity of heritage architecture.

2. IDENTIFYING CHANGES

2.1 Terminology

In order to avoid misunderstandings, it is important to define the two set of terms we will be using. To start with, we need to point out differences between the artefact as a whole and its possible remaining sub-parts once transformed (Table 1).

artefact	an entire object or ensemble considered as one basic entity
portion	a subset of an artefact resulting from its conceptual division into an active part (a core object) and an inactive part (a segment)
core object	an artefact’s main portion, *i.e.* - generally an apparent superstructure along with its active substructures, - possibly an inaccessible substructure when the apparent superstructure is demolished, - possibly, a substructure when the artefact was designed as a substructure from the start.
segment	a distinct, underground and inactive portion of an artefact

Table 1. Naming of the artefact and its sub-parts

We also need to make clear what we consider when talking about an *evolution of an artefact*, in comparison with what we call a *life cycle*.

- An ***evolution of an artefact*** is a process of gradual development of an artefact over a whole period of its life, from its creation until its extinction (*i.e.* its thorough and irreversible physical removal, including of sub-structures, or its division) or - if the artefacts still exists - until today.
- An ***artefact’s life cycle*** identifies a time slot corresponding to a fragment of artefact’s evolution. A time slot during which all transformations are partial (*i.e.* the artefact is neither moved nor entirely subdivided into new independent structures, its superstructure – if it had any - remains in elevation).

Given the above definitions, an artefact’s evolution may therefore contain several cycles of life. Good examples of this are Roman villas in the site of Pompeii, buried for centuries after the eruption of the Vesuvius. These villas have had (at least) three cycles of life: before the eruption, while buried (centuries below ashes), and since they have been uncovered by archaeologists and opened to visitors. Understandably, life cycles are then tagged as primal (first cycle after creation) or recurrent (others), simple or compound (if inactive substructures called segments exist underneath the artefact), and their sum marks the evolution of the artefact.

2.2 Two main notions: states and transitions

Each ***life cycle*** can consist of a number of ***states*** and ***transitions***. States are occupying time slots during which no major transformation occurs (or should we say when we have no indication that such transformations occurred). In other words, states identify periods of stability. Each state is preceded and concluded by transition(s) - time slots during which transformations occur. Often enough, transitions in the field of the architectural heritage may be rather long-lasting (*). Transitions indicate that a process of transformation is under way, with specific indications that underline possible causes (for instance, damages caused by fires, a common plague during the medieval period).

To sum it up, one can see transitions as causes, and states as consequences. Accordingly, their descriptions will naturally vary. In the next sub-section, we briefly introduce both states and transitions, and illustrate them by giving examples among which some are well-known artefacts or sites.

2.3 The tables of states and transitions

In this section we provide descriptions of states and transitions. Transitions and states may share the same tag: tags identify either the processes themselves (within which changes progressively occur, between a start date and an end date) or results of these processes (once changes denoted by the tag cease to occur). Typically, a decay transition starts when a first part of roof falls and ends when no roofing is left. The decay state then identifies the result of the decay transition: an artefact with no roofing, irresistibly degrading if nothing is done.

The tables contain a tag's name, its definition, and examples that should illustrate the semantics behind states and transitions (formatting presented below). Tags are ordered alphabetically.

tag	definition
example	
abandon <*transition*> [*state*]	Progressive withdrawing of a human activity, but artefact remains covered.

A medieval fortified village of Rougiers, erected on a hill 650 m over plateau, in 1760 became definitively abandoned by its inhabitants, who since XV century started to move to live on the plateau.

annexation <*transition*>	Combining/incorporating an artefact or its portion into another artefact. The life cycle of the annexed artefact (or portion) ends, while the annexing artefact continues the same life cycle.

Once the ensemble of Collegium Maius (Kraków) became a university (see transition "merge" below), it was further enlarged by acquiring new additional properties and integrating them into the existing ensemble. Annexed houses ceased to exist as independent artefacts, and became parts of the annexing ensemble.

decay <*transition*> [*state*]	A gradual degradation of an artefact due to absence of human activity and destruction of roofing.

It usually takes years for a building to loose its roofing (see for instance the Tivoli site).

demolition <*transition*>	A relatively quick destructive incident, caused by environment or human activity.

A Gothic town hall of Kraków was partially destroyed by fire in 1570 and 1680 and than renovated. In 1820 a main building of the town hall was demolished during a destruction of an adjacent building of granary. Only a tower and underground structures survived.

Kraków, town hall and granary before 1820

modification <*transition*>	Each significant functional, structural or morphological change inside one life cycle. This is naturally the most common transition.

The Louvre ensemble in Paris is the result of a number of modifications, either functional (at start- a defensive medieval castle, now- a museum) or morphological (medieval keep, classical and neo-classical galleries, contemporary sub-structures).

secession <*transition*>	A division of the artefact with a separation of one or several portions, leading to the creation of independent artefacts. The rest of artefact continues it's life inside the same life cycle.

A town hall of Saint-Maximin is a Neo-Classical edifice that belonged to Gothic monastic ensemble. In the beginning of XXth century it was separated from the monastery. Today it has nor functional neither morphological continuity with the rest of the monastic ensemble.

segmental anaesthesia <*transition*> [*state*]	Loss of all functional activity and of connexion with the rest of an artefact in one of its portions - without formal separation. It concerns only the underground structures

(a) Full-width perrons used to stand before urban houses around Kraków's main square, providing access to cellars and ground floor (1/2 level above ground). As in medieval period a level of the ground was quickly rising, perrons got concealed little by little and finally had to be walled in. They are today inaccessible (interred below the actual level of the main square)- (b). They are seen as segmental a naesthesia of the urban houses they belong to.

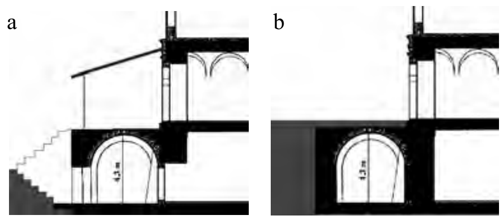


Table 2. Transitions and states occurring within a life cycle.

---

* It usually took several decades if not centuries to erect a gothic cathedral. When no interruption occurs during the building of an edifice, the whole period is seen as a transition. When on the contrary the work is halted for a significant period, for instance by wars, then a state is created that corresponds to an unfinished edifice. A number of examples exist, such as St Peter's cathedral in Beauvais, of edifices that in the end never were completed.

creation <transition> — Birth of an artefact (not to mix up with reincarnation, which is a re-birth).

Any act of creation of an artefact, even when long-lasting, provided that a continuous activity can be established on the work site (Sainte-Chapelle, 7 years).

extinction <transition> — Irreversible destruction of an artefact including sub-structures, with as result, its physical annihilation. Ends not only a life cycle but also the evolution of the artefact.

At the end of the Punic wars Carthago was intentionally destroyed by Romans, who as history tells us salted the area. Carthago was later replaced by a Roman town (later seized by Vandals, and finally destroyed during the Arab conquest) the remains of which one can today visit. The initial town is extinct.

hibernation <transition> [state]> — Destruction of the artefact's apparent parts, combined with absence of functional activity inside the remaining inaccessible portion.

Unexcavated edifices of the sites of Pompeii or Herculaneum that can be considered as hibernating until archaeologists uncover them.

internment <transition> [state] — Building of a new artefact over a previous one, the latter remaining underneath as an inactive, inaccessible portion called a segment. Internment may be deliberate (ex. preventive archaeological bury) or unintentional.

The Gothic Dominican monastic ensemble built in Saint-Maximin (to protect the relics of Sainte Marie-Madeleine whose tomb was discovered on the site in 1279) was erected on the remains of an antique Roman villa. The site was excavated in the past decades, but archaeologists and local authorities agreed on keeping the Roman remains (that also include an early baptistery) buried.

merge <transition> — Combining different artefacts or portions of artefacts into a union, that is considered as a completely new artefact.

Collegium Maius in Kraków (where Copernic studied) became the seat of Jagielonian University when king Casimir the Great bought from private owners various properties in order to create a new, compound yet consistent structure.

reincarnation <transition> [state] — Re-birth of an artefact in a new embodiment, occurring in particular as a result of excavation and reuse of a hibernating artefact.

The site of Knossos (Crete) was hidden (hibernating) for centuries before 8it was excavated and partly rebuilt at the end of the XIXth century, and then opened to visitors.

split <transition> — Formal division and separation of an entire artefact into parts, leading to the creation of two (or more) new, independent artefacts.

The former Palace of Diocletian, built by this Roman emperor in the Croatian city of Split, was not demolished but divided into a number of sub-parts, that today, after series of modifications are used as individual artefacts (urban houses in particular).

In some periods houses inside one urban block have q tendency to split (a result of proprieties division)

translocation <transition> [state] — Displacement of an artefact or sub-parts of an artefact, caused by human activity.

A part of a temple of Abou Simbel (Ramses II) was dismantled, moved and rebuilt higher backed on to an artificial cliff when lake Nasser was created.

Table 3. Transitions and states starting/ending a life cycle.

3. VISUAL DISPLAYS

3.1 Methodological background and graphic codes

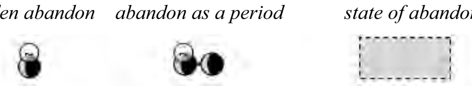
In the field of linguistics, F. de Saussure (see (Klinkenberg, 1996) or (Barthes, 1985)) identified three modes of time existence in human experience – *synchronia* (a time-slot in which an object of our analysis is the same, unchanged), *diachronia* (a period in which the object changes) and *panchronia* (what steps out of time-space continuum and authorise us to state the identity of the object regardless of changes). A thorough understanding of an artefact's evolution requires therefore three-mode analysis – synchronic, diachronic and panchronic.
The methodological approach presented in this paper proposes:

- a diachronic lecture of changes that allows recognition of transitions (sudden or gradual),
- identification of periods of synchronic existence that we call states,
- a pancronic understanding of evolution of an artefact, that allows us to affirm that regardless various changes the artefact rests the same, and state which transformations may alter artefact's identity.

The alternation of transitions and states, projected on the timescale gives us a coherent vision of artefact's evolution.
This framework is used to produce two types of linear diagrammatic representations called *diachrograms* and *variograms* (see Figure 2). The transitions and states are represented by distinctive graphic codes, that permit to read a diagram as a "story about an artefact's evolution".
All the basic graphic codes are represented in the table below. Note, that sometimes the difference between two distinct transitions is not directly expressed by a transition-code (e.g. annexation and merge), but it can be clearly identified by the consequences of a transition. (see Figure 1)

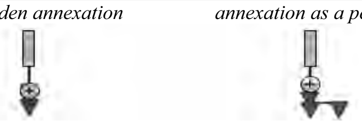
abandon *<transition>* *[state]*

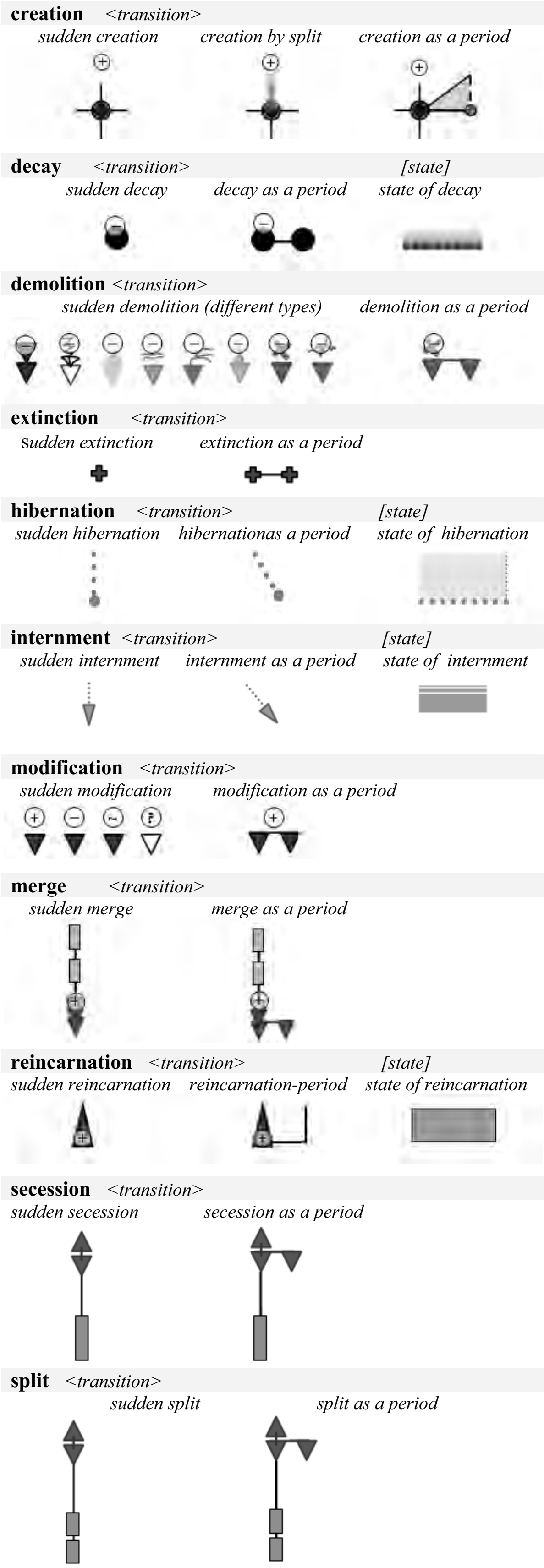
sudden abandon *abandon as a period* *state of abandon*



annexation *<transition>*

sudden annexation *annexation as a period*





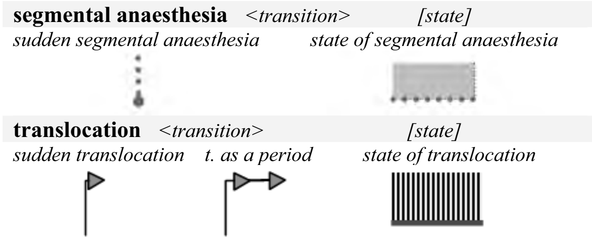


Table 4. Graphical symbols proposed to visualise identified transitions and states.

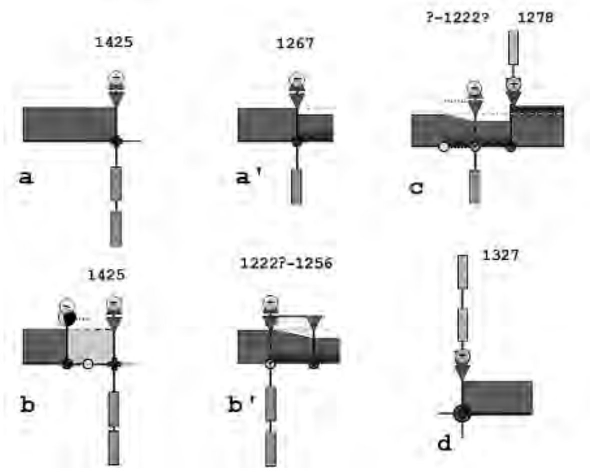


Figure 1: (a) sudden split transition of an artefact into two new independent artefacts, marked as two rectangles underneath the time axis; (b) same transition, after a period of abandon; (a') a sudden secession transition; (b') a lasting secession transition; (c) a lasting secession transition causes the artefact to diminish, and a following sudden annexation transition has the opposite effect (marked as a rectangle above the artefact's bar); (d) as a result of a lasting merge transition associating two artefacts, a new artefact is created;

3.2 Diachrogram

A *diachrogram* represents successive states and transitions for life cycles inside entire evolution of an artefact. They present the evolution of an artefact along a time axis.
The basic components of a diachrogram are a time axis, states and transitions markers, accompanied by the date-certitude indicators (see Fig. 2).

The date-certitude indicators informs us weather the dates of transitions are certain, dubious or vague. They are composed of a circle on the time axis and of a vertical line linking it with a corresponding transition marker. Colour of a circle (grey, dashed or white) and line type correspond to a certitude level.
One marker represents a sudden transition, two interrelated show a lasting transition (and its duration).
It has to be said that a diachrogram represent an expert's view of the artefact: different interpretation of the data may lead experts to propose different chronologies - the diachrogram then acts as a comparative tool.

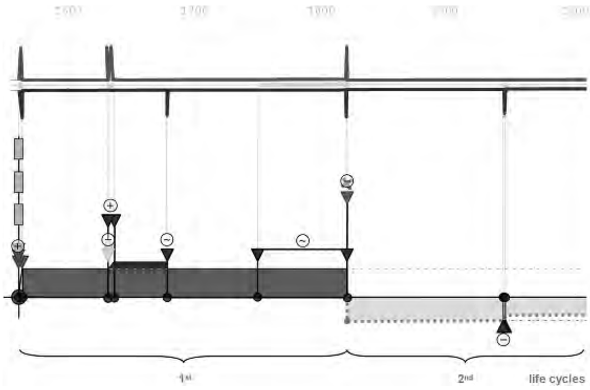


Figure 2: Combined diachrogram and variograms, with *(top)* morphological, structural and functional variograms representing the nature of changes, and with (*bottom*) the full diachrogram representing successive life cycles.

Diachrograms underline visually moments when the edifice tends to "get bigger" (as a result of modification by extension or of annexation for instance). They show the level of certainty or lacking information we have on the dating of events (start and end of events). (see Fig. 3) Colour coding is used to differentiate irregular states (abandon, hibernation, etc.), as shown in Table 4.

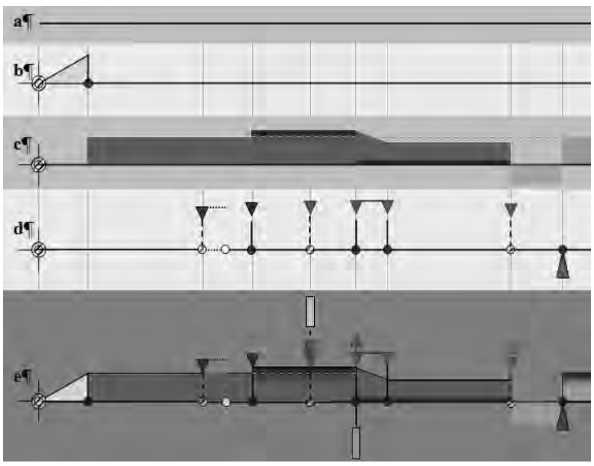


Figure 3: (**a**) the time axis; (**b**) symbols marking start and end of the artefact's evolution (creation) ; (**c**) the artefact's bar, composed of "states" with a dash/dot line acting as a reference to the artefact's initial volume; (**d**) transitions markers and date-certitude indicators (**e**); (**g**) a combination of components.

3.3 Variogram

A *variogram* visualises the nature and relative importance of transformations. It allows to highlight the intensity and duration of changes by combining in a parallel visualisation three aspects:

- morphological (formal changes such as stylistic refurbishing, changes in surface, volume, etc.),
- structural changes (technical changes such as change of roof material, replacement of sub-elements such as floors, etc.),
- functional changes (significant switches in the way the artefact is used).

3.4 Implementation

The implementation of the elements described in this paper, namely a framework of description of architectural changes, and visual disposals aimed at enhancing the readability and comprehensibility of the above mentioned changes, privileges open source and standards for the web. The developments presented in this paper in fact complement. previous works we have carried out on the same field of experimentation - the medieval heart of Kraków (presented for instance in (Blaise, 2007a) or (Blaise, 2005)) - with a focus put in these former actions on relations between 3D/2D graphics and a documentary database. Accordingly, the technical platform used is the same, and combines the following elements:

- a description of artefacts as instances of a hierarchy of classes (in the sense of OOP), with persistence enabled through RDBMS structures,
- outputs (may they be visual outputs – 3D VRML or 2D SVG- or textual outputs –XML) produced by Perl scripts nested in web-enabled pages,
- interfaces produced by Perl scripts either as XHTML or as XML/XSLT datasheets,
- graphics produced by Perl scripts either as VRML files or as SVG files (with included user-controls enabling various interactions within databases).

In this development, we have privileged dynamic SVG (see (Renolen,1997)); but at this stage SVG is used only for data visualisation, the interface itself remaining XML/XSLT (with embedded jscripts). The SVG format has been widely applied both in cartography (see *carto.net repository*) and in infovis. It appears as an efficient solution and allows the level of interaction we expect, with nested jscript when needed.

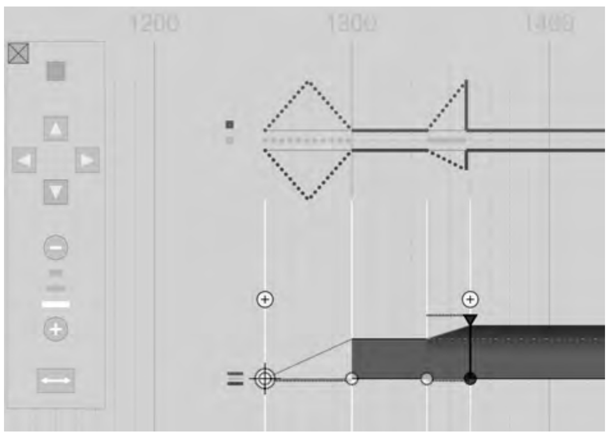


Figure 4: Partial view of a combined visualisation of diachrogram and variogram in the SVG implementation (here real case of the town's "Kramy Bogate", *i.e.* rich stalls). Navigation in and about the scene uses the classic geographic map browsing metaphor with plus and minus signs for zoom factor, and arrows for pan command.

Interaction inside the SVG display allows users not only movements inside a static scene but also control on various elements:

- graphic context (background time grid);
- date markers (white lines and background time grid can be used to check date of event);
- additional information (textual elements given as comment about the event).

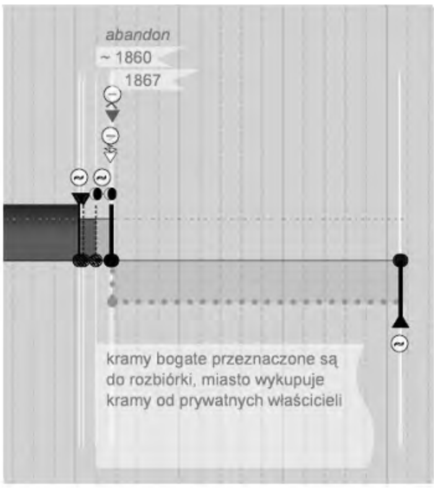


Figure 5: Additional information with specific uncertainty markers for dates (here real case of the town's "Kramy Bogate", *i.e.* rich stalls).

Both the evolution of architectural and urban elements (335 objects, 817 phases studied) and historical sources used during the investigation (761 sources) have been described. The separation of various types of data allows independent growth of each database or data set, and cumulating information brought by specialists from different domains.

Diachrograms and variograms will be produced dynamically for the above mentioned data set, with however a necessary evaluation period needed to test the disposal's efficency as "tool for thinking".

4. EVALUATION

At this stage of our research, evaluation of the disposal as a whole (both the description framework and the graphics) is done by confrontation with experts of the field of experimentation. This is due to the fact that the readability and efficiency of the graphics requires not only a good understanding of historic architecture, but a good knowledge of changes that specifically occurred in the city of Kraków throughout history (in order to point out lacks, underline inconsistencies or misleading visual signs, etc.). Figure 6 shows examples of applications to real cases in the medieval heart of Kraków that are being evaluated (Saint cross church, a middle size, minor religious artefact, and the old town hall, a civil, complex, major artefact).

Once this first evaluation will have been carried out, it may appear necessary to widen the audience in order to better evaluate to which extent the visual disposals themselves are efficient.

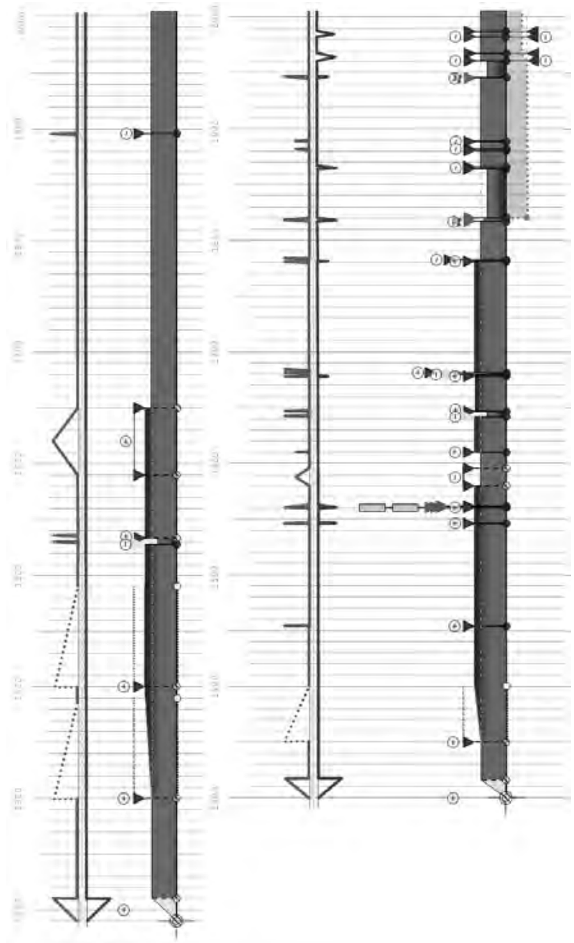


Figure.6: Top: Variograms and diachrogram for Saint Cross church: note absence of functional change (bottom line of the variogram), a typical feature of religious artefacts. Also note strong uncertainties on the dating of the two first transitions; and a short transition due to fire around 1530. Finally, note the stability since mid 17th century, corresponding to the period when the capital of Poland is transferred from Kraków to Warsaw (a drop in the city's wealth). Once cross-examined with buildings around, or with buildings of the same type, this might also indicate a lack of information. Bottom: Variograms and diachrogram for city's former town hall. Note the one-century time slot between the creation of Saint-Cross church and of the town hall. Also note the number of functional changes (bottom line of the variogram): the edifice was used as a town hall, but also during its evolution as a prison, a tavern, a granary, a museum, a theatre, etc.. Note numerous morphological changes, with two fires during the 17th century, and two partial demolitions (19th/20th century). Also, note consequences of the 19th century partial demolition: superstructures diminish; an inaccessible, inactive substructure appears (and remains as such until now). Finally, note the relatively good precision we have on the dating and duration of transitions (except for the first), due to the importance of the building, and therefore to the amount of documents we have on its evolution.

5. CONCLUSION

Observing that solutions lack when one wants to recount and sum up the evolution of historic artefacts (lacks in terms of method of description as well as of visualisation), we propose and apply a methodological framework dedicated at a diachronic reading of architectural changes. Graphics developed are primarily designed to allow the visual assessment of an artefact's life cycles. In addition, following E.R Tufte's observation (Tufte, 2001) - *comparisons must be enforced within the scope of the eyespan*- they should provide means for visual comparisons (time t_1 to t_n of an artefact's evolution, comparative reading of artefacts). We believe that the synthetic nature of these graphics helps understanding in a cleat-cut manner how changes over time affect architecture, but also underline key aspects of "historical sciences" data: uncertainties, incompleteness, long ranges of time, unevenly distributed physical and temporal stratigraphy, etc..

6. REFERENCES AND SELECTED BIBLIOGRAPHY:

Alkhoven, P., 1993. *The changing image of the city. A study of the transformation of the townscape using Computer assisted Design and visualisation technique*s. PHD Thesis, University of Utrecht.

Ando, M., 2003. Networked VR for virtual heritage. In : *HCI International 2003*, LEA publishers, pp. 1386-1390.

Barthes, R., 1985. *L'aventure sémiologique*. Seuil, Paris.

Blaise, J.Y, Dudek, I., 2005. From artefact representation to information visualisation: genesis of informative modelling. In : *SmartGraphics 2005*, LNCS, Vol. 3638, pp. 230-236.

Blaise, J.Y, Dudek, I., 2007. Informative Modelling, *MiaJournal,* Vol. 1S.

Blaise, J.Y, Dudek, I., 2007a. Visual tools decipher historic artefacts documentation. In: *JUCS I-Know 07 - 6th International Conference on Knowledge Management*, Gratz, pp. 456-463.

Bocheński, J.M.,1986. *The methods of contemporary thought*, Harper & Row, New York.

De Luca L, et al., 2005, Reverse engineering of architectural buildings based on a hybrid modelling approach. In : *Computers & Graphics*. Vol. 30

Estevez, D., 2001. *Dessin d'architecture et infographie*, CNRS Editions, Paris.

Huber, B., 2000. Adding dimensions to GIS with VRML. In : *Proc. Pennsylvania GIS Conf.*

Kantner, J., 2000. Realism vs Reality: creating virtual reconstructions of prehistoric architecture, In : J.A Barcelo, M.Forte, D.H Sanders (Eds.) *Virtual reality in archaeology*, Oxford: Archeopress.

Kienreich, W., 2006. Information and Knowledge Visualisation: an oblique view, *MiaJournal*, Vol. 0, pp. 7-16.

Klinkenberg, J.M., 1996. *General semiotics*. De Boeck & Larcier, Bruxelle.

Kodym, O., 1999. Usage of the VRML environment for the presentation of the underground mine. In: *Proceedings ICEE*.

Lenntorp, B., 2003, The drama of real-life in a time-geographic disguise. In : *Proc. 6th Théo Quant Meeting*.

Lerma, J.L, Vidal, J., Portalés, C., 2004. Three dimensional city model visualisation for real-time guided museum tours. *The photogrammetric record,* 19/108 .

MhMK 2007. Civitas Nostra Cracoviensis – A.D.1650 CD Rom published by the city of Cracow's Historical Museum.

Müller, U., et al., 1997. Damage mapping of historical buildings. In : *Proc SFIIC*.

Perkins, A., 2003. The Cone Sisters' Apartments: Creating A Real-Time, Interactive Virtual Tour, In : *Proc. ICHIM 03*.

Ramondino, F., 2001. From point cloud to surface : the modeling and visualisation problem, In : *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIV- 5/W10.

Rathert, N.A., 2005. Knowledge visualisation using dynamic SVG Charts, In: *Geroimenko, V., Chen C (Eds).: Visualizing information using SVG and X3D*, Springer.

Renolen, A., 1997. Conceptual modelling and spatiotemporal information systems : how to model the real world. In: *Proc ScanGIS 97*, pp. 35-53.

Sebillo, M., 2003. A Web-GIS for promoting archaeological assets. In : *Proc. ICHIM 03*.

Spence, R., 2001. *Information vizualisation*. Addison Wesley ACM Press, Essex.

Suveg, I., Vosselman, G., 2003. Reconstruction of 3D building models form aerial images and maps. In : *ISPRS Journal of Photogrammetry and remote sensing*, Vol, 58.

Tufte, E.R., 2001. *Envisioning Information*. Graphic Press, Cheshire.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

A 3D HUMAN ACTIVITY RECOGNITION SYSTEM FOR INTERACTIVE ENVIRONMENTS

Sujung Bae [a], Jin Choi [a], Yong-il Cho [a], Hyun S. Yang [a,]*

[a] AIM Lab., Computer Science Dept., KAIST, Daejeon, South Korea
(sjbae, jin_choi, caelus, hsyang)@paradise.kaist.ac.kr

KEY WORDS: 3D Human Activity Recognition, Interactive Environment, Motion Descriptor, Motion History Volume, View-invariant feature, Feature Dimension Reduction

ABSTRACT:

Activity recognition is an important and challenging topic as one way for human to interact with real or virtual objects in interactive environments. Especially 3D human activity recognition has recently begun to come into the spotlight since it is view-independent and robust to self occlusion which is the nature property that human motion has. This paper describes a 3D human activity recognition system which uses the approach proposed by Weinland et al., using motion history volume (MHV) as a motion descriptor and applying Fourier transformation to the MHV to have the descriptor view-invariant. In this paper, using the approach we tried to recognize not only the finished activity but also the activity being under way because it results in fast response which increases a feeling of satisfaction during interaction. Moreover, we made an effort to recognize interaction between people to show that it can be applied in various situations. To do that, we supplemented two additional features to original features. One is whether activity is upward or downward and the other is a radius of activity. For recognition on undergoing activity, we applied K-means clustering and seeking its nearest neighbour in classification step. Experiment results indicated that these efforts contribute to achieve our goal.

1. INTRODUCTION

1.1 Introduction

As growing information technology and computing resource, demands about smart environment get higher day after day. To meet the demand many researchers are struggle to develop such a system that can perceive the context and provide right service at a right moment. Activity recognition is an important and challenging topic as one way for human to interact with real or virtual objects in interactive environments. For the most part, recent researches about human activity recognition have tended to center around view-dependent representations. 3D human activity recognition has recently begun to come into the spotlight since it is view-independent and robust to self occlusion which is the nature property that human motion has.

This paper describes a 3D human activity recognition system which uses the approach proposed by Weinland et al., using motion history volume (MHV) as a motion descriptor and applying Fourier transformation to the MHV to have the descriptor view-invariant (Weinland et al., 2006). Using the approach we tried to recognize not only the finished activity but also the activity being under way because it results in fast response which increases a feeling of satisfaction during interaction. Moreover, we made an effort to recognize interaction between people to show that it can be applied in various situations.

To achieve our goal, we included additional features which are omitted by reducing original feature dimension. There are two additional features. One is z-orientation of activity which means whether activity is downward or upward. The other is a radius of activity measured from visual hull constructed from multiple silhouette images of object from different viewpoints. These features help raise feature description power to recognize various activities including people interaction. Also we improve on classification step by employ K-means clustering to recognize undergoing activity.

1.2 Related Works

Human activity recognition generally consists of 3 parts – pre-processing, modelling human activity, and classifying human activities. In pre-processing step, human movement, silhouette, and appearance are extracted using methods such as background subtraction (KaewTraKulPong & Bowden, 2001; Haritaoglu et al., 2000), skin detection (Jones & Rehg, 2002), and shadow removal(Prati et al., 2001). Data modelling process is process to make data which is later used to recognition process using information about human movement. There are two approach, structure information-based approach and global feature-based approach (Aggarwal & Cai, 1999). Structure information-based approach involves the low-level segmentation of human body into segments connected by joints and recovers the underlying 3D structure of the human body using its 2D projections over a sequence of images (Aggarwal & Cai, 1999). In this category there is two approaches, top-down and bottom-up approach. Top-down approach (Urtasun et al., 2005; Zhao et al., 2006) is that match a projection of the human body with the image observation and bottom-up approach (Ioffe & Forsyth, 2001; Ramanan & Forsyth, 2003; Ramanan et al., 2005) is that individual body parts are found and then assembled into a human body. Global feature-based considers whole body as a single unit and model and describe feature using appearance information such as shapes, contours, texture, silhouette, and motion extracted from image. Motion history image (MHI) (Bobick et al., 1997), Spatio Temporal

Volume Descriptor (Blank et al., 2005), and Action Sketch (Yilmaz & Shah, 2005) are methods in this category. This approach is accepted gladly since it is difficult to find correspondence of body parts between consecutive frames in spite of heuristic assumptions and predefined models in structure-based approach and appearance-based features relieve cost of computation.

For classification of human movement or activity, one effort has used state-space approaches to understand the human motion sequence. This approach defines each static posture as a state. These states are connected by certain probabilities. Motion sequence is a tour going through various states of these static poses. Joint probabilities are computed through these tours, and the maximum value is selected as the criterion for classifying motions (Aggarwal & Cai, 1999). In this technique, duration of motion is no more problem because state can be visited over again. However, this approach usually apply intrinsic nonlinear model which as Aggarwal and Cai said typically requires global optimum solution in the training process. It means expensive computing iterations. Finite state machine (FSM) (Bobick & Wilson, 1997), baysian network (Madabhushi & Aggarwal, 1999), and hidden markov model (Oliver et al., 2002) are in this approach. An alternative is to use the template matching. A model about human activity is converted into a static shape pattern and then it is compared to the pre-stored prototype patterns during recognition process. The advantage of this technique is its lower computational cost; however, it is usually more sensitive to the variance of the movement duration (Aggarwal & Cai, 1999).

Our previous work (Choi et al., 2007) adopted MHI, appearance-based human activity modelling method. Being common with other appearance-based methods, its features are view-dependant and weak to self-occlusion. To overcome these disadvantages, in this paper, we present a 3D human activity recognition system.

The paper is organized follows. First, we briefly introduce our system in Section 2. We present detailed method to model activity using MHV and extract view-invariant feature from it and classification in Section 3. Our improved method to recognize undergoing activities and interaction between people also described in Section 3. We show experiment results in Section 4 and conclude in Section 5.

2. SYSTEM OVERVIEW

Our aim is to provide a system which recognizes single person performed activities such as walking, running, sitting down, standing up, falling down, punching, kicking, and turning and interaction of two people such as hugging, and shaking. The system only uses sequence of images obtained from multiple cameras. Our system has to satisfy following constraints as much as possible.

2.1 System constraints

2.1.1 System can use 3D information about human body: It is natural for human being to construct and to use 3D information when they see and when they recognize objects in real world. That the machines use 3D information seems natural also. Usually 3D representation is more informative than 2D images so we can obtain more useful information.

2.1.2 System can be view-invariant: Until now, many human action recognition systems are view-specific so their results are dependent on orientation of viewpoint. Though changing viewpoint, recognition should be successful.

2.1.3 System can be independent on location and size of actors

2.1.4 System can operate in real-time: Many applications need real-time system. However, it requires high computation time to use multi-cameras and to satisfy above all constraints. When considering future usage in practical field, this constraint is considered.

2.2 System overview

Our proposed system is divided into four subtasks: finding silhouettes, modelling action, extracting motion descriptor, and classifying action. Detailed description of each task and its subtask is explained in later section. Figure 1 shows a structure of proposed system.

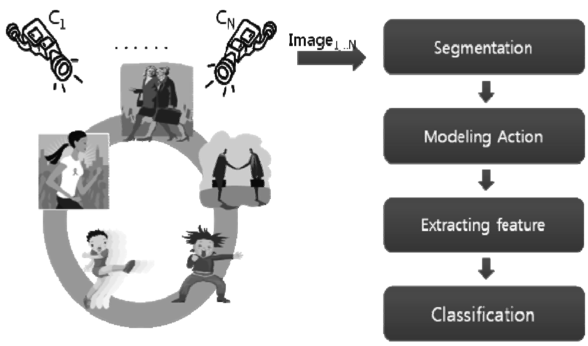


Figure 1: Overview of the system

3. 3D HUMAN ACTIVITY RECOGNITION

3.1 Segmentation

Our system needs to get rapid and correct silhouette from visual input since modelling action is based on the silhouette. Although there have been many algorithm to get silhouette, running average algorithm and mixture of Gaussian (MOG) are two popular background subtraction algorithms. Each of the two has disadvantage. Running Average algorithm can't cope with multimodal background distribution and standard MOG doesn't have that kind of problem but is too slow to be used in real-time system.

As Park reported in their paper (Park et al., 2006), the MOG algorithm with a hierarchical data structure yields results that are very similar to the results of the standard MOG algorithm and enhances the processing speed significantly so we can overcome standard MOG's week point. Algorithm of the MOG with a hierarchical data structure consists two parts: initialization and searching. Following is the algorithm.

3.1.1 The MOG algorithm with a hierarchical data structure: In initialization step, it requires constructing quad-tree based image region decomposition. Algorithm in searching step is presented by following box.

1 L = l_{init}
2 For each node in layer L
  2.1 A random pixel is sampled and classification of the pixel as foreground or background
  2.2 Update background model of that pixel location
  2.3 If the pixel is background then Continue
  2.4 If L != l_{final} then

It means that L saves layer number, l_{init} and l_{final} are set by user. Usually, for a 640 X 480 image, it is reported that $l_{init} = 4$ and $l_{final} = 6$ are adequate (Park et al., 2006).

As you can see above algorithm, however, it has still disadvantage that the user can't manage the form of the leaf nodes in this method so we use the method proposed by Choi et al. in our previous work (Choi et al., 2007). In Choi's method, which they call the method modified MOG algorithm with a hierarchical data structure, they constructs the hierarchical data structure using bottom up approach to manage the form to user's needs. When compared with Park's one, the only different thing is initialization part so we'll describe the part.

3.1.2 Modified MOG algorithm with a hierarchical data structure: The only thing different to previous algorithm is initialization step. In initialization step, a set of pixels are grouped into the form that we wish to detect from bottom layer. Then stack layers that consist of the parent nodes of four children (North West, North East, South West, and South East). After Stacking the layers to the designed layer, apply a quad-tree based decomposition to an image input (Choi et al., 2007). Figure 2 is an illustration of this approach in case of having three-layer hierarchy.

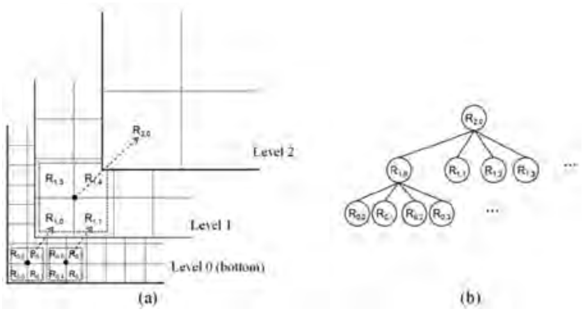


Figure 2: The three-layer hierarchical data structure formed with a bottom-up approach corresponding to image decomposition region (a). (b) is the built quad-tree. The structure is generated from level 0 to level 2 in order.

Figure 3: The results of background subtraction (the left image is an original image and the right image show the foreground image)

Obtained foreground pixels are grouped into a set of blobs using connected component analysis. This is for recognizing interaction of more than one people. As a result you can see Figure 4.

Figure 4: Input image and its result blob containing two peoples

3.2 Modelling Action

3.2.1 Visual Hull: Given multiple silhouette images of object from different viewpoints, its 3D shape can be constructed by intersecting the corresponding bounding volumes formed by the silhouette (Cheung et al., 2000).

In our system, however, we don't compute rays from silhouettes because Cheung et al. reported in their paper that intersecting rays in 3D is very numerically unstable and we think it seems tricky to implement. Instead we use robust 3D voxel reconstruction method proposed by Cheung et al.(Cheung et al., 2000). Although they didn't call the result obtained from their method the visual hull, we designate the result visual hull by Laurentini's definition (Laurentini, 1994) on visual hull which of a 3-D object S is the closest approximation of S that can be obtained with the volume intersection approach.

General 3D voxel reconstruction steps are as follows. The space of interest is divided into N X N X N equal sized voxels. Each voxel v is tested whether it belongs to foreground pixels by projecting itself onto all the k silhouette images. More formally, denote $Proj^k(v)$ as the projected region of voxel v onto k_{th} silhouette image and $Is\_Overlap^k(A)$ as function which return true when region A overlap with k_{th} foreground pixels or return false. Moreover INSIDE means a voxel belong to all k foreground object and OUTSIZE means it doesn't. With these definitions, voxel v is classified by following algorithm (Cheung et al., 2000).

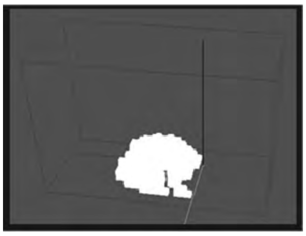
1 Set k, the index of silhouette image as 1
2 If $Is\_Overlap^k(Proj^k(v))$ is FALSE
  2.1 Then set v OUTSIZE, end
  2.2 Else if k is equal to K
    2.2.1 Then set v INSIDE, end
    2.2.2 Else set k = k + 1, goto Step 2

Problem of general approach is that it takes too much time to get visual hull since there are many pixels in the area generated by projecting voxel v and all these pixels should be tested whether these pixels are silhouette pixels or not. Cheung et al. proposed very simple algorithm to reduce computation time. They named it Sparse Pixel Occupancy Test (SPOT). They improved task of checking pixel's overlap with foreground object or not. Suppose on average S is the number of pixels inside the region generated by projecting voxel v. Instead of testing all the S pixels, they sample uniformly distributed Q pixels within the area and test the Q pixels. In this method, voxel v is belong to foreground objects if at least Q_ε samples ($\varepsilon < 1$) are classified to INSIDE pixels. SPOT is $\frac{S}{Q}$ times faster than general approach which requires N^3S tests. SPOT algorithm is stated formally as follows.

```
1     For voxel v and camera k, randomly choose Q
      points inside the projection of ∏^k(v),
      represented by q(l) where l = 1, … Q.
2     Initialize incount to 0.
3     For l = 1 to Q
      3.1      If q(l) is a silhouette pixel then
      3.2      incount = incount + 1
4     If incount < Q_ε then v ∈ OUTSIZE
      else v    INSIDE
```

In our system, we chose one point (Q = 1) located at the center of voxel since there remains many things to do after getting visual hull and we need faster reconstruction than SPOT. Figure 5 shows visual hull of walking people when we choose the center of voxel as testing point.

(a) Input images and their shilouette images



(b) A corresponding visual hull constructed from shilouette images

Figure 5: Visual hull of falling people

3.2.2 Motion History Volume: Motion history volume (MHV) is 3D motion templates based on visual hull proposed by Weinland, Rondfard, and Boyer to construct view-independent representation using calibrated cameras (Weinland et al., 2006). Because MHV is generalization of 2D motion templates, we shall start by an attempt to review 2D motion templates.

Bobick and Davis proposed motion energy images (MEI) and motion history images (MHI) as 2D motion templates (Bobick & Davis, 2001). MEI has information where there is motion therefore its pixel value is binary to capture motion occurrence at a pixel. MHI represents how recently motion occurs so its pixel value is multiple-value to encode it at a pixel. Let D(x, y, t) be a binary-valued function and D = 1 indicate motion at time t and location (x, y), then MHI is defined by

$$H_\tau(x,y,t) = \begin{cases} \tau & \text{if D(x,y,t) = 1} \\ \max(0, H_\tau(x,y,t-1)-1) & \text{otherwise} \end{cases} \quad (1)$$

where τ is the maximum duration a motion is stored. The result is a scalar-valued image where more recent pixels of motion are brighter. The MEI accompanied with MHI can be generated by thresholding the MHI above zero.

Now we'll examine 3D motion template proposed by Weinland, et al. 3D motion template is extension of 2D motion template with occupancy function D(x, y, z, t) and voxels instead of D(x, y, t) and pixels, where D = 1 if voxel in location (x, y, z) is occupied at time t and D = 0 otherwise. The more formal definition is like this.

$$H_\tau(x,y,z,t) = \begin{cases} \tau & \text{if D(x,y,z,t) = 1} \\ \max(0, H_\tau(x,y,z,t-1)-1) & \text{otherwise} \end{cases} \quad (2)$$

Bobick and Davis suggested substituting silhouette occupancy function for D(x, y, t) which capture only region where motion occur (Bobick & Davis, 2001). They argued that using whole body silhouette is more robust to incidental motions that occur during an action. This claim proved by Weinland, Rondfard, and Boyer in their experiments although including whole body may add another difficulty for recognition of motion since occupancy encodes shapes as well as motion (Weinland et al., 2006). According to suggestion of Bobick and Davis, the occupancy function D(x, y, z, t) is estimated using visual hull. Figure 6 shows an example of a MHV of a kicking motion. The first 4 images of Figure 6 show visual hulls made during kicking motion and the 5th image is the corresponding MHV. Figure 7 shows example of several MHVs.

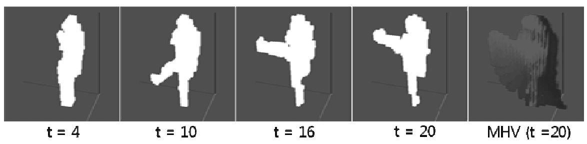
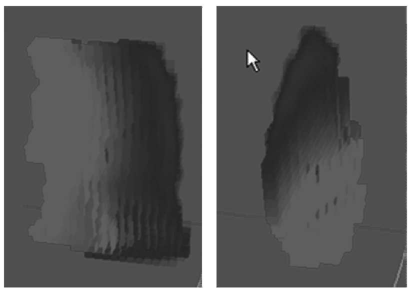
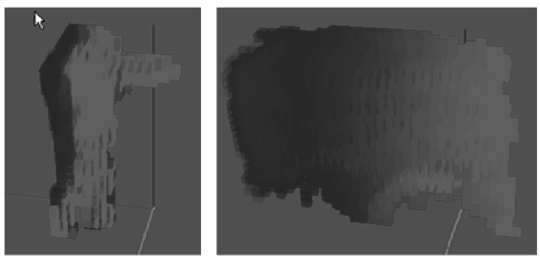


Figure 6: MHV of kicking, (a),(b),(c),(d) are visual hull during kicking motion, (e) is a corresponding MHV



(a) Walking (b) sitting



(c) punching (d) running

Figure 7: MHV examples

3.3 Extracting Features

After modelling human activity, making motion descriptor is important task since motion descriptor directly affects classification step. Motion Descriptors have to be free in human's location, size, and direction. Location and size dependencies are easy to remove than rotation dependency. Usual way to remove location and size dependencies are translating with respect to center of mass and scale normalization with respect to a unit variance.

Bobick and Davis decided to use Hu moments to extract features for MEI and MHI since Hu moments are known to yield reasonable shape discrimination in a translation- and scale-invariant manner (Bobick & Davis, 2001). Statistical models of the moments (mean and covariance) were generated for both the MEI and MHI and these were used in classification step. However, these features were still rotational dependent, they recorded same image from many angles as much as possible to prepare for changing view orientation in real environment. As it were, they avoided rotational dependency.

Weinland et al. made an experiment on descriptors extracted by using simple 3D extension of Hu mements (Sadjadi & Hall, 1980). They concluded that these descriptors are fail to discriminate detailed actions and not easy to use in practice in case of higher order moments. Also Shen and Ip showed these features are weak in noise (Shen, & Ip, 1999). Instead they suggested using Fourier based features. They observed Fourier based features are robust to noise and irregularities, and present the nice property to separate coarse global and fine local features in low and high frequency components. In addition, using Fast Fourier Transforms (FFT) Fourier based features can be efficiently computed (Weinland et al., 2006). The best thing is Fourier based features are view-invariant.

What makes Fourier based features rotation invariant? Weinland et al. paid attention to Fourier shift theorem: a function $f_0(x)$ and its translated counterpart $f_t(x) = f_0(x - x_0)$ only differ by a phase modulation after Fourier transformation (Weinland et al., 2006).

$$F_t(k) = F_0(k)e^{-j2\pi k x_0} \tag{3}$$

After removing phase modulation, the magnitude spectrum of $|F_t(k)|$ is given by

$$|F_t(k)| = |F_0(k)| \tag{4}$$

Therefore, Fourier magnitudes $|F_t(k)|$ become shift invariant representations. They thought by choosing proper coordinate system in which rotation in standard coordinate can be translation, it can be possible to make rotation invariant feature using Fourier shift theorem. They chose cylindrical coordinate system as proper coordinate system since rotation about x-y plane in standard coordinate can be rotation about θ in cylindrical coordinate system. They thought Fourier-magnitudes and cylindrical coordinate work well in most situations.

3.3.1 View-invariant Feature Vector: Now we explain the process to make view-invariant feature more specific. Our system passes through following steps.

The first step is translation normalization which is achieved by translating the origin to the center of mass of the MHV. The position of the center of mass (x,y) is determined by following.

$$x = \frac{\sum_{i=1}^{N} m_i x_i}{M}, \quad y = \frac{\sum_{i=1}^{N} m_i y_i}{M} \tag{5}$$

where M equals total mass and we assume $m_i = 1$ so M equals N. The second step is changing the MHV in a cylindrical coordinate system. Let v(x, y, z) be an intensity function of voxel in location (x, y, z). Then its cylindrical coordinate representation is

$$v(\sqrt{x^2+y^2}, \tan^{-1}\frac{y}{x}, z) = v(r,\theta,z) \tag{6}$$

As we mentioned above, rotation on the z-axis presents cyclical translation about θ in a cylindrical coordinates.

$$v(x\cos\theta_0 + y\sin\theta_0, -x\sin\theta_0 + y\cos\theta_0, z) = v(r, \theta + \theta_0, z) \tag{7}$$

The third step is scale normalization which makes variance σ r and σ z in z and r direction equal to 1. Next step is 1D Fourier transformation. To make feature view invariant, 1D Fourier transformation is required. A feature vector is composed of intensity values corresponding for every theta for selected (r, z) pair. The 1D Fourier transform for the feature vector is

$$V(r,k_\theta,z) = \int_{-\pi}^{\pi} v(r,\theta,z)e^{-j2\pi k_\theta \theta} d\theta \tag{8}$$

After 1D Fourier transformation for the feature vector, the Fourier magnitudes of the vector are independent of rotation along θ. An important property of the 1D Fourier transform is its trivial ambiguity with respect to the reversal of the signal. Consequently, motions that are symmetric to the z-axis result in the same motion descriptors (Weinland et al., 2006). For example, motion descriptors of people raising a left arm and people raising a right arm are same. In some application, this feature may be considered as loss of information; however, we think it is not positively necessary to discriminate between motions symmetric to z-axis in recognizing simple human activity.

3.3.2 Feature Dimension Reduction: We make a final 1D feature vector by concatenating Fourier magnitudes one by one for all (r, z) pairs. In this case, dimension of feature vector are very high so it burden computational cost to our system. Therefore we need its dimension to be reduced.

PCA is common method for dimensional reduction. Data points are projected onto a subspace that is chosen to yield the

reconstruction with minimum squared error. It has been shown that this subspace is spanned by the K largest eigenvectors of the data's covariance ∑, and corresponds to the directions of maximum variance within the data (Weinland et al., 2006).
For further data reduction, we performed linear discriminant analysis (LDA) by reason that we can't be convinced that K largest eigenvectors which PCA finds are useful for discriminating between data in different classes. Discriminant analysis finds directions that are efficient for discrimination while PCA finds directions that are efficient for representation (Duda et al., 2001).

3.3.3 Additional Features: Not only reducing feature dimension dramatically but also keeping feature's discriminant power high, we used PCA and LDA. But we can't avoid losing useful information to discriminate classes none the less. Therefore we decided to supplement two features after LDA that are Z-orientation of activity and a radius of activity.
Z-orientation of activity means whether activity is downward or upward. This information use to distinguish various activities. The activities affected from the information are jumping, sitting, standing, and ascending the stairs, etc. To express z-orientation, we used two constant values, ß1 and ß2, which are chosen from min value to max value of the feature vector elements after applying LDA for all training samples.
A radius of activity is added to help recognize interaction between people because some MHV generated from interaction between people is very similar to other MHV generated from activity performed by a single person. For example, walking and hugging. It helps not only distinguish interaction between people from single person performed activity, but also discriminate among activities by single person. The more activity shows broad movement toward x or y direction, the bigger the radius is. For example, running and walking have large radius value while sitting, standing, and turning have small one. The value is measured from visual hull. To normalize this value, we sought min of these values from training data. And we divide the radius measured from visual hull by the min.

3.4 Classification

To get high recognition rate, we have to consider few difficulties in classification step. First thing is the way people perform same activity is very different. Moreover, though a person repeats the same activity, they sometimes acts in the different way. Second, recognizing not only the finished activity but also the activity being under way raises more difficulty. For example, the MHV from turning activity and the MHV from standing almost done are very similar. Third, as we mentioned already, recognizing two kinds of activities, activity by one person and interaction between people, also weight extra confusion. Considering these things, we conclude that K-means clustering and seeking its nearest neighbour (NN) is proper method. K-means algorithm is composed of the following steps (Hartigan & Wong, 1979).

1. Select K points on data. These points represent centroids of initial group.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the positions of centroids no longer are changed.

After finishing K-means clustering for each class, each class is represented by K mean vectors. Any unknown sample z is then classified by finding its nearest mean vector (Simard & Mitran, 1999).

$$\begin{aligned} &\min\ d(m_i^k, z) \\ &d(m_i^k, z) = \left\| m_i^k - z \right\| \end{aligned} \tag{9}$$

where i = 1,…,the number of activity and k = 1,…,K. In section 4, experiments show you its propriety.

4. EXPERIMENTAL RESULTS

4.1 Experimental Environment Settings

We used 4 cameras and saved movie clip about activities of twelve subjects from each camera. 4 silhouette images are extracted from images achieved from each camera and they are used to carve visual hull. 20 frames including 19 past frames and current frame are used to build MHV at each time moment. After the MHV mapping into cylindrical coordinates, the representation has a resolution 24 × 24 × 24. The representation through 1D –Fourier transformation now has a resolution 24 × 24 × 6 since we ignore high frequency.
For training, we asked 7 subjects to conduct the ten distinct actions of walking, running, sitting down, standing up, falling down, punching, kicking, turning, hugging, and shacking two times. Since we didn't have any method about temporal segmentation, the beginning and the ending of activities was set manually. Thus the durations of activities are different for each person. We achieved 327 largest eigenvectors from training data and using these we reduced 1D-Fourier magnitudes to 327 dimensions. For LDA, we adopt direct LDA (D-LDA) which is applicable to small sample size problems (Wu et al., 2004) to reduce 327 dimensions to 8 dimensions. And finally two additional features are added to this. Now our training data has 10 dimensions.

4.2 Experimental Results

For experiments, 7 subjects performed same kind of activities as training. 2 subjects among 7 subjects have participated in training. Feature vectors were obtained the same way as training. Each subject performed same activity twice. For measurement, we manually set starting frame and ending frame for every activity of each subject and measure recognition rate between the starting and ending frame. We recorded average duration per activity.

Activity	Average duration (frames)
Walking	33
Running	15
Sitting	27
Standing	14
Falling	35
Punching	17
Kicking	20
Turning	20
Hugging	21
Shaking	37

Table 1: Average duration per activity

The first experiment is to show how much additional feature raise discriminant power. We compared result from 1 mean including additional features with 1 mean not including additional features. As you can see in Table 2, two additional features raise discriminant power by 0.86%. Especially these features affected single person performed activities.

Activity	Average Recognition Rate (%) Without additional features	Average Recognition Rate (%) With additional features
Walking	95.80	96.27
Running	100.00	100.00
Sitting	96.93	97.32
Standing	67.91	77.01
Falling	88.65	89.11
Punching	92.30	92.30
Kicking	95.68	95.68
Turning	95.88	96.47
Hugging	81.52	79.03
Shaking	94.38	94.38
Average rate (%)	90.90	91.76

Table 2: Average recognition rate applying 1 mean with and without additional features

The second experiment is to show how much K-means contribute to recognition rates. In Table 3, we compared result from 1 mean with 8 means. We didn't apply additional features in this experiment. As you can see in Table 3, K-means is a better solution than 1 mean to recognize undergoing activities since it can catch various states happening under activity.

Activity	Average Recognition Rate (%) with 1 means	Average Recognition Rate (%) with 8 means
Walking	95.80	94.75
Running	100.00	100.00
Sitting	96.93	94.94
Standing	67.91	76.58
Falling	88.65	94.97
Punching	92.30	99.61
Kicking	95.68	96.83
Turning	95.88	92.86
Hugging	81.52	86.53
Shaking	94.38	97.87
Average rate (%)	90.90	93.49

Table 3: Average recognition rate comparing result of 1 mean with K-means

Finally, there is result showing excellence of our method, a combination of K-means plus two additional features in Table 4. While we recorded average recognition rate is 90.90% when applying one mean using original feature, the combination recorded average recognition rate is 95.57%. Our method improves the overall classification results by 4.67% compared with original one.

Activity	Average Recognition Rate (%) with original method	Average Recognition Rate (%) with improved method
Walking	95.80	96.00
Running	100.00	100.00
Sitting	96.93	98.51
Standing	67.91	93.68
Falling	88.65	94.97
Punching	92.30	98.52
Kicking	95.68	96.83
Turning	95.88	92.78
Hugging	81.52	87.57
Shaking	94.38	96.85
Average rate (%)	90.90	95.57

Table 4: Average recognition rate comparing result of 1 mean without 2 additional features with 8 means with 2 additional features

5. CONCLUSIONS

We adopted MHV and method of extracting view-invariant feature from the MHV proposed by Weinland et al. and applied PCA and LDA to reduce feature dimension. Especially, LDA allow our feature dimension to be drastically reduced to 8 components. However, as LDA was doing this, important components to discriminate classes are wiped out. To supplement this, we include two features, z-orientation and a radius of activity. These additional features raise recognition rate of both activities, a single person performed activity and interaction between people. Moreover, to obtain speedy mutual response during interaction, we tried to recognize activities being under way. We observed that K-means give better result compared with using 1 mean in this case through experiment. There was a synergy between applying K-means and including 2 additional features so it raised overall classification rate.
We manually set starting frame and ending frame since we ignored temporal segmentation about activities. To enhance our system, automatic temporal segmentation should be considered. We observed variation of movement duration affect recognition. Developing less sensitive feature of the variation is also our future work.

References

Aggarwal, J.K. and Cai, Q., 1999. Human Motion Analysis: A Review, *Computer Vision and Image Understanding*, 73(3), pp. 428-440.

Blank, M., Gorelick, L., Shechtman, E., Irani, M., Basri, R., 2005. Actions as Space-Time Shapes, *International Conference on Computer Vision*, 2, pp. 1395- 1402.

Bobick, A.F. and Davis, J.W., 1997. Action Recognition Using Temporal Templates, Motion-Based Recognition, *Computational Imaging and Vision*, pp.125-146.

Bobick, A.F. and Davis, J.W.,, 2001. The Recognition of Human Movement Using Temporal Templates, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3), pp. 257-265.

Bobick, A.F. and Wilson, A.D., 1997. A state-based approach to the representation and recognition of gesture, *Pattern Analysis and Machine Intelligence*, 19(12), pp. 1325-1337.

Cheung, K.M., Kanade, T., Bouguet, J., 2000. Holler, M., A Real Time System for Robust 3D Voxel Reconstruction of Human Motions, *IEEE Conference on Computer Vision and Pattern Recognition*, 2, pp. 714 – 720.

Choi, J., Cho, Y., Han, T., Yang, Hyun S., 2007. A View-based Real-time Human Action Recognition System as an Interface for Human Computer Interaction, *International Conference on Virtual Systems and Multimedia*

Duda, Richard O., Hart, Peter E., Stork, David G., 2001. *Pattern Classification*, Wiley-Interscience, pp. 114-121.

Haritaoglu, I., Harwood D., Davis LS, 2000. W4: Real-Time Surveillance of People and Their Activities, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 73(3), pp. 809-830.

Hartigan, J.A. and Wong, M.A., 1979. Algorithm AS 136: A k-means clustering algorithm, *Applied Statistics*, 28, pp. 100-108.

Ioffe S. and Forsyth D., 2001. Human tracking with mixtures of trees, *International Conference on Computer Vision*, 1, pp. 690-695.

Jones, M.J. and Rehg, J.M., 2002. Statistical Color Models with Application to Skin Detection, *International Journal of Computer Vision*, 46(1), pp. 81-96.

KaewTraKulPong and P., Bowden, R., 2001. An improved adaptive background mixture model for real-time tracking with shadow detection, *2dn European Workshop on Advanced Video Based Surveillance Systems*

Laurentini, Aldo, 1994. The Visual Hull Concept for Silhouette-Based Image Understanding, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2), pp. 150-162.

Madabhushi, A., Aggarwal, J.K., 1999. A bayesian approach to human activity recognition, *IEEE Workshop on Visual Surveillance*, pp. 25-32.

Oliver, N., Horvitz, E., Garg, A., 2002. Layered representations for human activity recognition, *IEEE International Conference on Multimodal Interfaces*, pp. 3- 8.

Park, J., Tabb, A., Kak, A. C., 2006. Hierarchical Data Structure for Real-Time Background Subtraction, *IEEE International Conference on Image Processing*

Prati, A., Cucchiara, R., Mikic, I., Trivedi, M.M., 2001. Analysis and Detection of Shadows in Video Streams: A Comparative Evaluation, *Computer Vision and Pattern Recognition*, 2, pp. 571-576.

Ramanan, D. and Forsyth, D.A., 2003. Finding and tracking people from the bottom up, *Computer Vision and Pattern Recognition*, 2, pp.467-474.

Ramanan, D., Forsyth, D.A., Zisserman, A., 2005. Strike a pose: Tracking people by finding stylized poses, *IEEE Conf. on Computer Vision and Pattern Recognition*, 1, pp. 271-278.

Sadjadi, F. and Hall, E., 1980. Three-Dimensional Shape Analysis Using Moments and Fourier Descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(2), pp. 127-136.

Shen, D., Ip, H. H.-S., 1999. Discriminative wavelet shape descriptors for recognition of 2-D patterns, *Pattern Recognition*, 32(2), pp. 151-165.

Simard, P. and Mitran, M., 1999. A Tutorial on course “Pattern Recognition“, Nearest Neighbor Rule: A Short Tutorial, http://cgm.cs.mcgill.ca/~soss/cs644/projects/simard/(accessed 19 Sep. 2008)

Urtasun, R., Fleet, D.J., Fua, P., 2005. Monocular 3–D Tracking of the Golf Swing, *IEEE Conference on Computer Vision and Pattern Recognition*, 2, pp. 932-938.

Weinland, D., Ronfard, R., Boyer, E., 2006. Free Viewpoint Action Recognition using Motion History Volumes, *Computer Vision and Image Understanding*, 104(2-3), pp. 249-257.

Wu, X., Kittler, J., Yang, J., etc., 2004. A new direct LDA (DLDA) algorithm for Feature Extraction in Face Recognition, *International Conference on Pattern Recognition*

Yilmaz, A. and Shah, M., 2005. Actions sketch: a novel action representation, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1, pp. 984- 989.

Zhao, Q., Kang, J., Tao, H., Hua, W., 2006. Part Based Human Tracking In A Multiple Cues Fusion Framework, *International Conference on Pattern Recognition*, 1, pp. 450-455.

Acknowledgements

This research is supported by Foundation of ubiquitous computing and networking project (UCN) Project, the Ministry of Knowledge Economy(MKE) 21st Century Frontier R&D Program in Korea and a result of subproject UCN 08B3-O4-10M.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

ENVISION OF A DIGITAL FOUNDATION TO EMBED KNOWLEDGE AROUND RELEVANT ASPECTS OF CULTURAL HERITAGE

Giuseppina Enrica Cinque [a], Luigi Mazzucchelli [b]

[a] University of Rome “Tor Vergata”, Engineering Faculty, 00100 Rome, Italy – cinque@uniroma2.it
[b] NEXT-Ingegneria dei Sistemi SpA, Research and Innovation, 00100 Rome, Italy – luigi.mazzucchelli@next.it

KEY WORDS: Digital Spaces, Digital Maps, Representation of Knowledge, LBS, Identification, Unified approach, 3D representation

ABSTRACT:

Reviewing current results and applications in the domain of (digital) cultural heritage, such as virtual reality, digital information management, multimedia, makes evident that technologies and systems are nowadays pervasive in the investigation phases and in the analysis and exploitation of the scientific data. Applications of virtual reality by means of 3D representations of architectural and historical models are attracting a wide range of audience – especially in the tourist and multimedia sector. Typically, the emphasis is given to the visual representation and its impact to the user, giving less importance to the scientific accuracy and capacity of the platform to represent the whole “knowledge” associated to the cultural asset.

There is a demanding need to envision approaches to communize all the elements of such multidisciplinary knowledge. A data set of coherent, structured, multidisciplinary, multi-source, multi-media information (that can be physically distributed) - constituted by “markers” (i.e. pertinent to a specific domain such as architectural, technical, structural, historical, bibliographical, archive related, photographical, physical, chemical, ...) – should give the capacity to implement Cultural Heritage platforms around a sort of a “Digital Foundation” of assets to support different applications and user experiences. Current approaches and availability of digital libraries, the accessibility of distributed information (e.g. by internet), the capacity to model data and meta-data, the availability of instruments to accurately position objects in space and time (together with ontological representations) let us envision a novel model of representing a dynamical evolving multi-source “digital foundation”.

The above model, as the result of a multidisciplinary approach coming from a long experience in research and technologies in various fields (ICT architectures, Digital Libraries, Knowledge representation, Distributed infrastructures, GIS and applications of Satellite Navigation) and projects (National and European) is now experimented around the UNESCO Villa Adriana in Tivoli (Rome, Italy). It constitutes the center of excellence that allowed the authors to preliminary build a digital foundation for the area, demonstrating the capability to generate (among other features) accurate 3D virtual reality representations, distributed multi-source/multi-target and multi-modal representation for virtual guides and to support archeological studies and excavations for professional users. It furthermore implements capacities to manage, to control and to develop the necessary knowledge for the safeguard and conservation of the immense material associated to archeological sites.

1. INTRODUCTION

Current results and applications in the domain of (digital) cultural heritage, such as virtual reality, digital information management, multimedia, makes evident that technologies and systems are nowadays pervasive in the investigation phases and in the analysis and exploitation of data.

Applications of virtual reality by means of 3D representations of architectural and historical models are attracting a wide range of audience – especially in the tourist and multimedia sector. Typically, the emphasis is given to the visual representation and its impact to the user, giving less importance to the scientific accuracy and capacity of the platform to represent the whole “knowledge” associated to the cultural asset.

There is therefore a demanding need to envision approaches to communize all the elements of such multidisciplinary knowledge. A data set of coherent, structured, multidisciplinary, multi-source, multi-media information (that can be physically distributed) - constituted by “markers” (i.e. pertinent to a specific domain such as architectural, technical, structural, historical, bibliographical, archive related, photographical, physical, chemical, ...) – should give the capacity to implement Cultural Heritage platforms around a sort of a “Digital Foundation” of assets to support different applications and user experiences. Current approaches and availability of digital libraries, the accessibility of distributed information (e.g. by internet), the capacity to model data and meta-data, the availability of instruments to accurately position objects in space and time (together with ontological representations) let us envision a novel model of representing a dynamical evolving multi-source “digital foundation” called “Digital Space”.

The above model, as the result of a multidisciplinary approach coming from a long experience in research and technologies in various fields (ICT architectures, Digital Libraries, Knowledge representation, Distributed infrastructures, GIS and applications of Satellite Navigation) and projects (National and European) is now experimented around the UNESCO Villa Adriana in Tivoli (Rome, Italy). It constitutes the center of excellence that allowed the authors to preliminary build a digital foundation for the area, demonstrating the capability to generate (among other features) accurate 3D virtual reality representations, distributed multi-source/multi-target and multi-modal representation for virtual guides and to support archaeological studies and excavations for professional users. It furthermore implements capacities to manage, to control and to develop the necessary knowledge for the safeguard and conservation of the immense material associated to archaeological sites.

The approach elucidated in the present work was experimented in the course of European Project called CUSPIS in the framework of applications of Satellite Navigation for the Cultural Heritage sector.

2. VILLA ADRIANA AS A CENTER OF EXCELLENCE

2.1 The UNESCO Site of Villa Adriana

Figure 1: Villa Adriana, Canopo

Hadrian's Villa - **Villa Adriana**- close to Tivoli (Rome, Italy) - is one of the most important and well known monumental complexes of Roman antiquity. It is among the most visited cities and museums in Italy, reaching a maximum presence of about 400.000 visitors per year. The Villa was built by the emperor Publius Elius Adrianus between 117 and 138 C.E. In 1999 the archaeological area of the villa was acknowledged by UNESCO as one of the Human Heritage Monuments. According to UNESCO, it can be considered the symbol of the Roman Empire since it is considered among the highest forms of ancient culture and one of the main models for the study of the birth and development of modern architecture. This unique character is also emphasised by the many exhibitions that are dedicated to Adriano every year around the world. This uniqueness was clearly wanted by Hadrian for his Villa, since it was meant to be the place for personal leisure. This character expresses also his own cultural approach, as seen in other massive buildings in Rome (The Temple of Venus, The Mausoleum known today as Castel Sant'Angelo, Elio bridge and the Pantheon) and in the Roman provinces (Hadrian Wall, The Arches of Gerasa and Athen).

Such an important characteristic is immediately clear in all the structures of Villa Adriana. The Villa is a complex of building sprawled over some 120 hectares and consisting of several structures (some 28) each one variously shaped which can be singled out even if they belong to a global project, probably created in part by Emperor Hadrian himself (117-138 C.E.). The vastness of the project can be better understood from a comparison with Pompei, that covers an area of about 60 hectares only. Today only 770 rooms are open to visitors and these are connected by open spaces, gardens, and criptoporticoes and, especially by an extraordinary underground road network originally dedicated to pedestrians and carts and not yet completely known.

Villa Adriana has been excavated for more than five centuries and is famous worldwide, but it lives in a paradox since it is still scarcely known in its essence and structure – as many other Roman monuments. The reason for this partial and fragmented knowledge of the Villa can be traced to the nature of old excavations and studies, often limited to an antiquarian and humanistic perspective: most of them dealt with just one single building, without taking into consideration the Villa as a whole. Ancient sources, from Renaissance on, simply copied from each other the antiquarian information, and some were simple fairytales about its treasures. There are few comprehensive studies on Villa Adriana as a complex, sometimes outdated or incomplete and often providing limited and insufficient information. One major reason is the extreme difficulty encountered to gain a global perspective of each single structure, mainly due to progressive decay, restoration measures and excavations. Most of the research completed so far has been focused on archaeological aspects of single components or artefacts, without any general framework that could encompass the information that was progressively unveiled.

As an example, it shall be mentioned that most of the available surveys are affected by errors, and that they represent only a superficial image of the real geometry of many of the existing structures. The exact position of the different structures within the complex is presently unknown. At the same time, only a minor part of the entire complex is known in detail. Furthermore, the location and the general layout of the hydraulic and transportation systems are only in part known, even if this knowledge would be of great help when proposing hypotheses with respect to the original design of the Villa. Finally it should be noted that many structures are still buried in the ground; should they be unveiled, the information on their characteristics would add significantly to the general knowledge of the complex. This would certainly help in promoting the value of the Villa with respect to the management of tourist flow and in general its sustainability.

The need for a comprehensive study that stems from a multidisciplinary survey is strong. This comprehensive study would form a general framework and integrated layout to link new specialised information related to structural, architectural and physical aspects that could be unveiled using innovative techniques. Based on such framework, as an example, it would be possible to develop an economical model to manage tourist access and maximise a sustainable use of the Villa with great cultural benefits.

This aim is strongly supported in Italy by the **Soprintendenza ai Beni Archeologici del Lazio**, that has recently signed an agreement with University of Rome "Tor Vergata" in order to develop a modern tool to store survey data in a GIS form and that is supporting a whole set of projects around Villa Adriana to form a "center of excellence" to develop innovative projects at all levels - including a new technological perspective.

Specific activities have been defined in order to verify and check all the information that has been retrieved on the Villa in the past and that will be progressively unveiled in the future. This includes data on scientific areas that are quite different and even far apart from the fields pertaining to archaeology. Hence, the foundation will form a new starting basis, updated and corrected in future studies.

2.2 Studies cantered around Villa Adriana: the universal map

Several studies were developed under the request of the Soprintendenza per i Beni Archeologici del Lazio to obtain a comprehensive survey and therefore a sort of "universal" geo-referenced 3D map of the entire land upon which is built Villa Adriana. Such a "universal map" should then constitute a basis for all further developments of studies and investigations for scientific, archeological, historical and other studies around the area. In this respect, all management of the interventions, maintenance, restoration and even – tourist visits – should accomplish to use and rely on such a map. Villa Adriana major features, to include, among others, construction materials, decorations, hydraulic functions, soil-structure interaction, stability, resistance, typological solutions, construction details, were thoroughly investigated. In this respect it shall be pointed out that at the present state of knowledge Hadrian's Villa resembles a model of intense architectonical experiments. Different typological solutions have been adopted after several functional modifications; constructed volumes have been organised and arranged in such a way to form interacting and complicated spaces. The engineering aspects relate to the vast excavations aimed at reshaping the local terrain, to the specific solutions adopted to cover several complexes, to the extended network of underground constructions, dedicated to pedestrian and carriage pathways as well as to the numerous hydraulic conduits. It shall also be pointed out that the hypotheses and related model are compared with existing structures still surviving within other archaeological complex that have been traced to Adriano throughout the Mediterranean. In essence the models and the hypotheses pertaining to structures of the Hadrian's Villa that are no longer existent will be compared and validate against structures of the same period traced to Adriano.

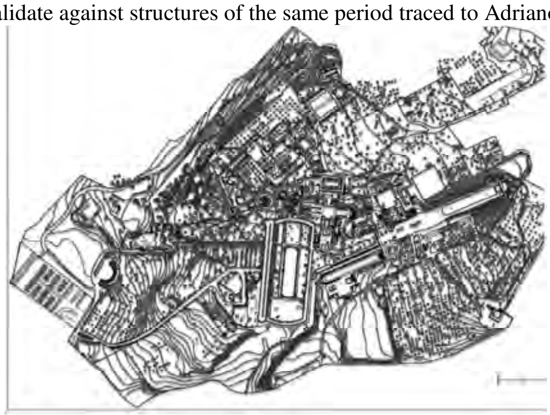


Figure 2: 2D projection of the Villa Adriana "Universal Map"

The planimetric complexity of the area, the existence of a rich set of underground features (including roads), the diverse attitude of buildings and houses with respect to their functional features, the complex shape of the terrain and the different states of conservation and their positional and geometric attributes requested numerous surveys with different and international experts. Different and up-to-date methodological and technological processes were applied to build the universal map – including:

- A completely new approach to the digital library archive structured in layers and linked to GIS facilities;
- The use of aerial photogrammetry, 3D scanners, reference stations:
- The use of suitable User Terminals to encompass GPS/EGNOS and storage models studies appropriately for the scope;

To the scope of the correct interpolation, extraction and reconstruction of geometrical features, link to the correct data stored in the digital libraries and data-bases and the need to use data for 2D/3D visualization and reconstruction lead the team (coordinated by the University of Rome Tor Vergata) to study new approaches for the cataloguing, storage and representation of the data.

In this respect, the need to comprehend data in all forms (geographical, geometrical, historical, reconstructive, …) is enclosing the whole background of knowledge around pieces, parts, whole of assets in the area. Such a representation constitutes a core to:

- Represent objects in diverse dimensional features;
- Allow dynamical reconstruction and representation of objects in diverse contexts and towards different users (tourist, archeologists, managers, institutions,…);
- Contains all relevant features to allow 2D/3D representation and management of data (e.g. in digital libraries).

As an example, a reference case-study is targeted to build the digital representation of a part of Villa Adriana –the building of Roccabruna – as a reference example to enclose 3D reconstruction in different time-frames, accurate modeling of parts and the whole building with accurate 3D scanner surveys and the successive intervention of the experts to models complex spaces and data.

In this respect, the approach pursued here is distinguishing for the following features:

- Easy extraction of features in both space and time frames;
- Accurate geometrical representation and compatibility with most common tools for rendering (incl. Revit);
- Capability to link objects with external databases and digital libraries (by means of standard identification and tagging);
- Capability for the model to be appropriate in different contexts (i.e. tourist, archeologist, …) by means of suitable meta-data layers;

Such a model encompasses usual 3D representation in both terms of quality as well as for the different perspective: it constitutes a way to construct the knowledge of the asset and not its representation.

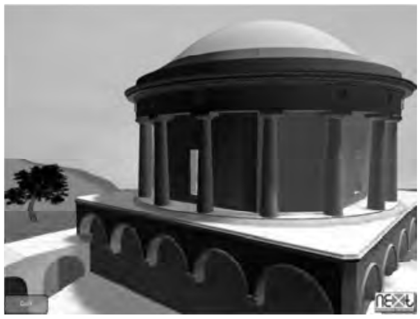


Figure 3: 3D representation of Roccabruna at Villa Adriana

2.3 From visualisation to representation

Digital reconstructed models (2D/3D) of the various constructions will be first arranged in a specific catalogue and will provide a base for further developments to embrace its **Digital Space foundation**. The following objectives will be pursued:

1) Structuring of the (digital) library and the information representation in the fields pertaining to the following areas: historical documentation, census and typological classification of structures (with particular reference to the vaults and the domes), characteristics of decay and damages, degree of structural safety. The inventory will be, afterwards, functional to the realization of the "universal map".

2) Draft of a "map of the risks" in which the results of the comparison between survey data, analyses on static and

characterization of the parameters and building materials, executed with technologies of pertinence of the unity of nuclear physics, will converge. In particular the territory of the Villa will be assigned different risk levels and priorities of safeguarding measures will be assessed together with related guidelines for consolidation and restoration; a careful comparison between technical data and cultural, social and economic evaluations will be carried out to supply required suggestions and indications to accomplish the last objective.

3) Identification of "fruition models" and "plans for economic improvement" that consider all the potentials of the Villa. Such a plan shall stem from and combine available experiences related to the “theme parks " and "open air museum structures"; it shall allow to verify the sustainability (economic, financial, institutional), in relation to a cultural project able to involve not only the occasional visitor, but the entire user community.

The possibility to attain the proposed objectives is depending on the methodological and procedural coordination of the synergy from different scientific expertise, on the possibility to direct them towards the expected results and to results which are more favourable from a cultural and economical perspective in the area of ancient monuments. Finally, all the expected results should be included in a web accessible portal since the beginning of the first investigations, in order to allow the many existing users in the world to keep up to date information in real time.

In this respect it is worth remembering that, as an example, many parts of the Villa Adriana are still unknown and that several others have never been carefully investigated; also knowledge of a very high percentage of the complex is spoiled by restorations which have not been fully documented or are described in dispersed and fragmentary documents. With particular reference to the retrieval of archive data, a fundamental part of the work will be developed together with personnel of the Authorities that will be involved in the research; such work will be carried out in addition to the acquisition of information in other archives and libraries.

Also, the collection of data in situ can be organized to obtain information without heavily interfering with the administrative and tourist daily activities. Therefore techniques of indirect survey are best suited to obtain initially:

- Precise and certified geographic positioning of all structures included in public property areas and identifying of a realistic extent at a given time.
- Location of underground structures (or still uncovered).
- Geo-referenced location of every complex in relationship to the general map of the area.
- Realising supporting topographic survey and photogrammetry of visible and underground structures.
- Identification of assets to enable authentication and certification of objects.
- Classification and use of digital information towards standardised identification of objects and their relation to Digital Libraries and/or data bases.

Test of the methodology focuses on qualitative control of digital data obtained by the use of GPS/EGNOS[1] where Information will be processed through specific software and will be arranged and used to help develop an IT platform for management and (graphic) representation of all available data. This software is basically constituted by a Location Based - LBS - platform extended by the interoperability with Digital libraries tools.

Expected results of the construction phase can be divided in two specific interacting branches; the first concerns the upgrading of knowledge of the area to obtain:

- three-dimensional graphic restitution of survey in CAD environment
- Location of the areas with underground structures.
- Zoning of open and restricted areas in the villa.
- Classification of the images and the alphanumeric data in a relational database.
- Design of geographic database and some implementation procedures in GIS environment.
- The appointment of the necessary LBS platform composed of User Terminals (usually PDAs) equipped with GPS/EGNOS and communication to a Service Control Center (as the front end to the IT back-end system).

The following results are expected instead from a second group of activities, addressed to the qualitative control of techniques for data collection:

- Refining of technical procedures for survey data exchange.
- Alignment of geometrical data entries by heterogeneous techniques.
- Functional validation of the informative system architecture.

The second phase will be essentially completed in the laboratory and targeted at the realisation of the "universal map" for the archaeological area, being a product where finding objective information about sector analysis and researches. Information will be arranged to achieve the following objectives:

- Building typologies classification
- Building phases determination
- Diachronically identifying of the structures
- Analysis of building techniques and use of materials
- Typological classification of static behaviour of the structures
- Test of nuclear techniques in dating and characterizing materials
- Analysis of structural decay characteristics: fissures, collapses, detachments of pictures
- Classification and physical chemical characterization of materials
- Location and classification of materials decay phenomena
- Identification of areas to be valorised

Resulting data will allow also managing and programming all the excavations, conservation, restoration and tourist flows (and activities).

Data included in the "universal map" will be functional to the development of researches and to improve available knowledge of the area for diverse applications (including geo-archaeology, management and tourist related services).

[1] EGNOS is the Satellite Augmentation System managed by the European Space Agency to complement GPS signals in order to extend the accuracy of position and to provide initial services (such as integrity) in view of GALILEO.

3. THE USE OF LOCATION BASED SERVICES

3.1 Background on Location Based Services

Location-based services are any service that takes into account the geographic location of an entity. Even though location-based services (LBS) were predicted to become the "killer application" of mobile commerce, their dominance has not yet materialized— but is predicted to do so soon. The LBS market size has been predicted to grow exponentially from 2006 to 2010. Within this four-year time span, for example, Asia's LBS market is expected to increase from $291.7 million to $447 million, Europe's market from $191 million to $622 million, and the U.S. market from $150 million to $3.1 billion. Generally, the slow adoption of LBS has been explained by the following causes.

First, the implementation of accurate localization techniques through providers has taken longer and has been more costly than expected –in this respect and demonstrated by the experiences of several solutions in the Cultural Fruition services, EGNOS adoption, as a **Global Navigation Satellite Systems** (GNSS) European solution to support the United States GPS, will play the expected differentiator role. One of the major challenges of the European GNSS is in effect to develop new markets for services and products, using EGNOS as precursor of Galileo. Second, the few available LBS applications display difficulty to adapt to user willingness and behavioral orientation that mostly depends on human factors that are not fully taken into account by the actual systems (mostly with predetermined contents and without any model of behavior instilled in the system). And third, users are concerned about privacy issues that are an inevitable side effect of LBS.

A LBS platform is usually composed of a **Service and Control Center level**, a **User Terminal level**, a **Service** and a **Content level**. The following are examples of services enabled by LBS in the Cultural Heritage domain.

3.2 Cultural Fruition Services (CFS)

The cultural tour/visiting guide are supported by accurate positioning service allows receiving Location Based cultural data content visiting city historical centres or archaeological areas. The cultural information should be geo-referenced and customized for the user; the data retrieval process should take into account the proximity to a cultural asset (every cultural asset is associated to its spatial coordinates associated to other attributes in a point of interest), the user profiling information and the user preferences. The tourist should experience an interactive multimedia cultural content in which the right information is accessed at the right time and in the right place. The cultural asset fruition scenario is basically composed of two capabilities:

- The outdoor (open sky) visit (such as archaeological areas with located assets such as statues and complexes).
- The indoor visit (museums, art galleries).

These capabilities are distinguished from different factors:

- <u>The satellite signal availability</u>. In fact the satellite signal is not available in indoor environments. In consequence of that the user localization could be done with different technologies (such as RFID or Wireless).
- <u>Different communications mean availability</u>. In large outdoor environments is not usually expected to have broadband communication mean such as WiFi which is easily possible to have in indoor environments.
- <u>Different concept of maps</u>. In the outdoor environment is expected to have a territorial map (in vector or raster form). In the indoor environment is expected to have a building map composed by different floors and topological structure.

3.3 Geo-Asset management services (GAMS)

Geo-asset management is the concept of identifying, collecting, cataloguing and disseminating geo-tagged information for assets. This approach, seen in cooperation with Geo time authentication (**GTA**) services using GNSS, should allow archaeologists and expert users to use geo-related information about assets for excavations, studies and research (in view of **geo-Archaeology** and **landscape-archaeology**).

Basis of this approach is the availability of an "**identity**" of an asset. A cultural asset identifies a work of art (e.g. sculpture, picture, etc) that is catalogued into a system. Each cultural asset should associate with a unique ID. Such unique code (to be defined at national or international level) is usually assigned by public organizations.

This unique listing code can be furthermore associated to the GNSS signals. The longitude and latitude coordinates (i.e. the geographic coordinates) are the ones provided by GNSS to identify the place where the cultural asset has been catalogued. A GNSS timestamp could be then added to build a unique identifier for each cultural asset for a specific location. Suppose that two cultural assets are catalogued in the same place (i.e. they have the same coordinates). Then the timestamp generated by the **GNSS** system will allow discriminating among them. Such "certified" coordinates (as enabled by EGNOS and in future GALILEO, and not enabled by GPS) can additionally provide a mean to address the problem of a global standard that permits to catalogue different cultural assets residing in different places and to "certify" the authenticity of the asset itself. The additional use of physical tags (e.g. RFIDs) can be also support in tangible identification also for management purposes.

4. A ENVISION OF A DIGITAL FOUNDATION: A PERSPECTIVE

To enable a unified approach to different applications and platforms – based on a homogenised representation of assets - including geo-graphical attributes, data and its representation (in space and time) lead us to introduce the novel definition of a "**Digital Space"** as a foundation for representing archaeological areas and its constituents objects.

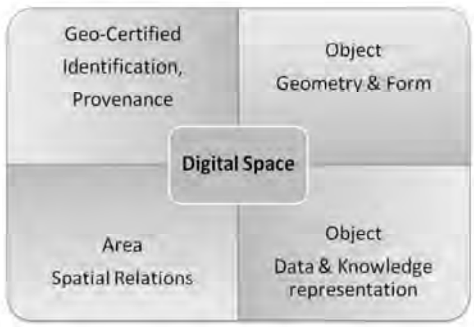


Figure 4: The Digital Space foundation constructs

This digital representation is an innovative instrument since it would constitute the basis for any informative system to supplement the data represented in diverse dimensional forms:

- The **physical form** – will be identified univocally by the **Global Asset Identifier** – complemented by the geo-position (particularly suited for open archaeological areas). This identification can be used for authentication, referencing and provenance purposes and can be easily stored in RFID tags.

- The **virtual form** (to be visualised in both in 2D or 3D forms) can be identified, referenced and stored as a digital data object. It has to be noted that the virtual representation objects can be:
    - Representation of areas with its relevant spatial structure of containment (e.g. a statue within a temple);
    - Representation of objects within a time-frame;
    - Geometrical representation;
    - The link to the actual instance in a defined digital library;
    - The description of the object (in form of data and meta-data);
    - The actual piece of art;

- The **digital library representation** in order to obtain the necessary platform for storage, handling and fruition. Within a coherent framework, the informative system will contain all the data already available in the existing literature and also those one that will be unveiled in different areas of study during the progress of the research.

- The **Spatial structure of areas,** to include relationships among objects in physical spaces (i.e. areas) to help digital representations to be located, related in space and topologically assembled.

The outline of the process to unveil digital space data can be then described with respect to investigation areas and steps to incrementally form "knowledge" of the area and objects contained in it:

- Survey and graphical restitution via 2D/3D dimensional models, area and point models (including Point of Interest POI) and relations between entities on the ground and virtual space-time reconstructions
- Storing of data and of information pertaining the fields of history and provenance, archive, references; classification and analyses of the structures existing above and below ground, with respect to their use and essential characteristics; classification, analyses and survey of decorations along walls and floors; classification and analyses of the hydraulic and transportation infrastructures
- Construction of the digital space by linking the different dimensional levels and the data base; this representation can be interrogated and navigated at different level on interest, in order to allow further analyses and also to arrange for different types of information to different types of fruition (i.e. tourist/experts) – maintaining the invariance of the supporting IT infrastructure

In essence, this methodology will provide a basis to convey the different scientific and methodological contributions and expertise towards the common objective of the construction of Digital Spaces.

Appropriate IT platforms should be appointed to let the digital construction be successful and usable to users and experts[2]. In fact, assets/objects can used to drive actions of the supporting platforms using relationships, proper ontologies and concepts structured around the "Digital Space Foundation".

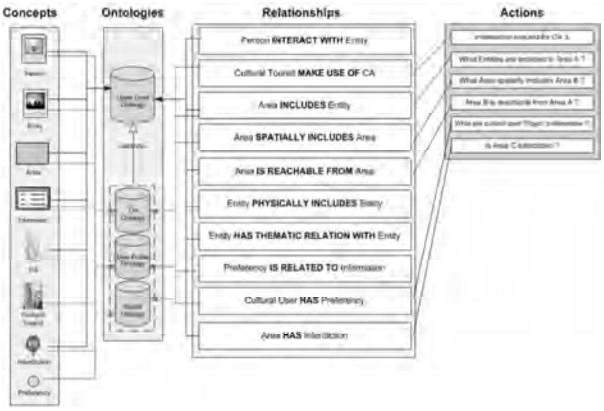


Figure 5: The Digital Space supporting capabilities

4.1 Authentication of assets and the Global Asset Identifier

The idea of authentication is based on the availability of a mean for authentication, as sort of electronic signature on the signal itself that can enable a number of applications and cases where the "trustworthy" and "authentication" of the location and time is considered an added value.

The authentication is based on the following elements:

A) **GEO-TIME authentication and tagging equipment/s**

These equipments are to be used on surveys to record geo-time data that are considered important to "certify" and authenticate assets (i.e. constituent objects in a physical space). A receiver (usually a PDA equipped with GPS/EGNOS for geo-tagging and communication device to link to external data bases) is needed. This receiver will store a local database of the recognised geographical areas linked to other specific data such as text/picture or video.

B) **GEO-TIME authentication tag**

The authentication tag could be then encapsulated/embedded in the asset (with a combination of visual readable information for customers and RFID based information for automatic check). The visible readable information is easy to be forged by malicious actors while the RFID tag could store cryptographic information. Such tag constitutes both the Geo time Identifier as well as the storage of provenance data.

Actually, a geo-time stamp is composed of the following fields (as did for the Villa Adriana assets in the universal map):

1. **User**: identifies the user that generated the ID;
2. **Starting time (i.e. year-month-day / hour-minute-seconds)**: defines the time in which the ID validity is activated;

[2] See the CUSPIS project as a reference platform in this respect.

3. **End time (i.e. year-month-day / hour-minute-seconds)**: defines the time after which the ID validity will expire. Time frame can be unlimited;
4. **Identifier:** identifies the asset ID in relation to a specific library and/or a standard (e.g. using DOI – the Digital Object Identifier).
5. **Type:** identifies a museum, an exhibition, maps service information, routing service information and points of interest service, ...;
6. **Area:** Identifier, in its topological structure, of the area (or sub-area) where the timestamp is valid.

The literature points out the need of developing an international standard to identify assets (ASSET ID or more recently the DOI structure), moreover it suggests the GNSS technology use as an added value to address the problem. A large number of cultural assets share the problem of identification over the world. The proposed unique marking (as the Global Asset ID) is indeed crucial to cataloguing and protecting a cultural asset so that public institutions and stakeholders can manage assets and objects in complex areas. The GAID can be associated to new objects found in a site both in archaeological events and when new discoveries are unexpectedly made (road constructions and so on). The GTA system can support these archaeological discovers and geo-referencing for mapping in order to uniquely identify every new CA and for precise and accurate positioning. Moreover, the GNSS signal can enhance security services in all data related applications for:

- **Authentication**: the ability to grant that the subject X is the subject X without any doubt.
- **Confidentially**: the ability to grant that what subject A is saying to subject B is not listened from an unwanted subject C.
- **Integrity**: the ability to grant that what subject B is listening from subject A is exactly what subject A is saying to subject B.
- **Non-repudiation**: the ability to grant that what is confirmed from subject A cannot be undone.

<u>Usage of Cultural Asset Identification within Digital Spaces</u>

The Digital-Space enabled platform should allow identifying **univocally** a cultural asset by:

- Assigning **GAID** (Global Asset Identifier) with the complement of **geo-time stamps** (**hash functions** are used to generate univocal keys)
- **Certifying Identity** by an Institution/authority using appropriate mechanisms and process related platforms. The certified data are saved in the Authority database and stored on the CA RFID.
  Support the verification of the **authenticity** of assets reading the (RFID) identity and checking it though the application and the certification authority

Additionally:

- The Digital-Space enabled platform application can generate "certified" **POI** from asset data, linking geo-time to identification and enabling IPR management (e.g. using GAID for watermarking);
- Institutions can issue **certificates** by using global GAID and geo-time data (certificates can be used in various applications such lending, insurance and transportation). Secure functionalities are therefore based on **certified data** (e.g. using **Crypto-keys**);

<u>Spatial data structures for Digital Spaces</u>

The spatial and cultural asset data must be organized in an abstract conceptual model to allow the semantic organization of spatial and relational concepts. The first step is to discretize the environment in areas essentially by means of two kinds of relationships – the Spatial inclusion and the Reachability

The reachability relation is an equivalence relation, in fact:

- Area A is reachable from area A.
- If area A is reachable from area B then area B is reachable from area A.
- If area A is reachable from area B and area B is reachable from area C than area A is reachable from area C.

The spatial inclusion relation is based on geometric considerations in fact the area A is spatially included in area B if the bounds of area B contains the bounds of area A. It is possible to see that this relation is asymmetric, in fact if area A is included in area B then are B cannot be included in area A. Each area may contain several entities, each entity is directly linked with:

- Its relational attributes
- Its spatial representation as a point (Point Of interest).
- The children entities: the entities which are physically contained in the father entity

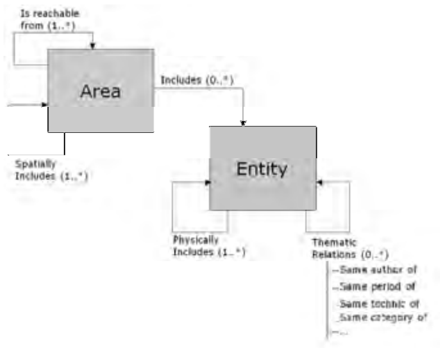


Figure 6: The Digital Space spatial data structure

Each entity could be in relation with other entities for physical inclusion or for other thematic relationship. In the same way the areas could be in relation for spatial inclusion or reachability. These relationships go to create a graph based structure that will be use to manage all relations and containments in the archaeological areas.

<u>Including Provenance data</u>

The concept of *Provenance* is already well understood in the study of fine art where it refers to the trusted, documented history of some work of art. Given that documented history, the object attains an authority that allows scholars to understand and appreciate its importance and context relative to other works of art. Objects that do not have a trusted, proven history may be treated with some skepticism by those that study and view them. This same concept of *Provenance* may also be applied to data and information generated within a computer system; particularly when the information is subject to regulatory control over an extended period of time. Today's digital libraries architectures suffer from limitations, such as lack of mechanisms to trace results and infrastructures to build up trusted networks. *Provenance* enables users to trace how a particular result has been arrived at by identifying the individual and aggregated services that produced a particular output. This support includes a scalable and secure architecture, an open proposal for standardizing the protocols and data structures, a set of tools for configuring and using the provenance architecture, an open source reference implementation, and a deployment and validation in industrial context.

The impact of the use of Provenance aware infrastructures (such as the one developed by open-source by the FP6 GRID Provenance project) is to provide mechanisms that allow information generated and managed within a grid infrastructure to be proven and trusted. By this we mean that the information's history, including the processes that created and modified it, are documented in a way that can be inspected, validated and reasoned about by authorized users that need to ensure information controls have not been altered, abused or tampered with.

5. RESULTS AND EXPERIMENTATION OF THE CONCEPTS

5.1 The CUSPIS VIth FP project results

In the context of the 6th Framework Programme and under the developments of the GALILEO calls, the Galileo Joint Undertaking launched a specific call reserved to the Special User Communities (Call 2-1A) to facilitate the introduction of GNSS and in particular of EGNOS and Galileo in the community of Cultural Heritage. The project **CUSPIS – Cultural Assets Space Identification System** (see http://www.cuspis-project.info) was a response to the call with a Consortium that comprehended all actors in the value chain and whose objectives were the following:

- Establish and focus on proper User Community Groups in the Cultural Asset context, enhancing their awareness of the potentiality of the Satellite Navigation toward the problems related to the Cultural Heritage domain;
- Analyse Cultural Assets issues related mainly to security, management and fruition;
- Identify all potential applications enabled by GALILEO services in the Cultural Asset sector;
- Analyse the potential applications, examining all the elements that can affect the final services; this activity will involve the key actors of the value chain, including the User Community related ones, and will assess the associated market potentiality. Throughout a trade-off analysis, some of the identified application were selected to be implemented and demonstrated,
- Perform a business and cost benefit analysis,

Although the project were focused on the User Community groups, fostering on analysis and penetration activities (in a field where navigation and communication technologies are of a novel application), a set of proof-of-concepts were demonstrated successfully to the Users with a character of innovation. The realisation of the CUSPIS platform that deployed the Authentication and Identification services (based on the concept of GAID), the demonstration of the geo-archaeology, management and fruition services at Villa Adriana allowed the authors to experiment successfully the concept of Digital Spaces in concrete applications. CUSPIS were disseminated in several events both at National as well at International level, demonstrating the innovative concepts behind the peculiar ideas envisaged in the design of mobile Cultural Fruition and the innovative models behind the Cultural Heritage domain.

CUSPIS studied and implemented a set of applicative services as a first experiment to exploit the concept of "Digital Spaces". Further experimentation is now to be launched to extend the work around Villa Adriana and also to export the model in other important excavations such as Pergamon in Turkey.

6. CONCLUSIONS AND FURTHER WORK

The considerations above and the promising results of successful implementations part of FP6 projects such as CUSPIS – together with the studies around Villa Adriana, managed and supported by key institutions, universities and industries, have helped to advance at both scientific and technological levels in several areas with a strong accent on validation of key technologies for the use and structure of digital spaces to enable a technology-enhanced learning and a knowledge construction process called "Digital Space".

The capability to experiment and advance on developed technologies and/or open-source tools constituted a key success factor for the present work and in-kind value to construct a reference platform for Cultural Heritage scenario as experimented around Villa Adriana.

7. REFERENCES

- Abruzzese D., Cinque G. E., Lo Gatto G., Analysis of a Roman masonry flat-slab in Hadrian's Villa, Tivoli, in Structural Analysis of Historical Constructions, Proc. Structural Analysis of Historical Constructions Conference, Padova, 2004.
- Adembri B., Cinque G. E., Villa Adriana, la Pianta del Centenario, Firenze 2006
- Adembri B., Archeologia on-line: Archeoguida di Villa Adriana, in MIBAC in rete, Roma 2006, pp. 52, 54
- Casella G, Tomei A., Torcinaro M., Il complesso di Roccabruna della Villa Adriana di Tivoli in Gambardella C, Martusciello S, a cura di, le vie dei mercanti-disegno come topologia della mente, atti del terzo forum internazionale di studi, Firenze 2006 pp 113-123
- Cinque G. E., Modelli sperimentali per la fruizione turistica della Villa Adriana di Tivoli, in Pasqualini, A., a cura di, Il Turismo culturale in Europa, Roma 2005 (2)
- Chiappetta F., I percorsi antichi di Villa Adriana, Roma 2008
- Mac Donald W.L., Pinto L.A., Hadrian's Villa and its legacy, Yale University 1995
- "CUSPIS: A reference Platform for LBS in Cultural Heritage Domain", Brussels Growing Galileo Conference 2008
- D. Pletinckx Ed. The Integration of Location Based Services in Tourism and Cultural Heritage", 2007
- Proceeding of VSMM 2007 (Part A and B)
- VAST2007 – Future Technologies to Empower Heritage Professionals
- http://www.doi.org. The DOI reference model
- http://www.esa.int . The GALILEO and EGNOS programs
- http://www.cuspis-project.info/ . The CUSPIS project website

8. ACKNOWLEDGEMENTS

The authors would like to thanks the Soprintendenza per i Beni Archeologici del Lazio and Benedetta Adembri for the important support and grant to the studies. This paper was partially funded by the 6th FP European Project CUSPIS.

Graphics Applications and Visualization Techniques

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

TECHNIQUES FOR THE INTERACTIVE EXPLORATION OF HIGH-DETAIL 3D BUILDING RECONSTRUCTIONS USING THE EXAMPLE OF ROMAN COLOGNE

S. Maass a, *, M. Trapp a, J.E. Kyprianidis a, J. Döllner a, M. Eichhorn b, R. Pokorski b, J. Bäuerlein c, H. v. Hesberg c

[a] Dep. of Computer Graphics, University of Potsdam, HPI, Germany -
(maass, trapp, kyprianidis, doellner)@hpi.uni-potsdam.de
[b] KISD, University of Applied Sciences Cologne, Germany - (michael.eichhorn , rafael.pokorski)@kisd.de
[c] Archaeological Institute, University of Cologne, Germany - j.baeuerlein@uni-koeln.de, hesberg@rom.dainst.org

KEY WORDS: High-detail 3D Models, Virtual Reality, Real-Time 3D Visualization, Roman Cologne

ABSTRACT:

This paper presents the results achieved by an interdisciplinary team of archaeologists, designers, and computer graphics engineers with the aim to virtually reconstruct an interactive high-detail 3D city model of Roman Cologne. We describe a content creation pipeline established to enable a flexible exchange and enhancement of building models, the applied optimization techniques necessary for real-time rendering, and the design of an application framework that enables the coupling of 3D visualizations with additional information in corresponding Adobe® Flash® widgets. Furthermore, we expose challenges arising by incorporating state-of-the-art visualization techniques, such as cut-away views, non-photorealistic rendering (NPR), and automated label placement. These techniques are used to enhance the interactive 3D environments, to enable for the presentation of interior structures, the precise communication what is hypothetic and what proven knowledge, and the integration of meta-information.

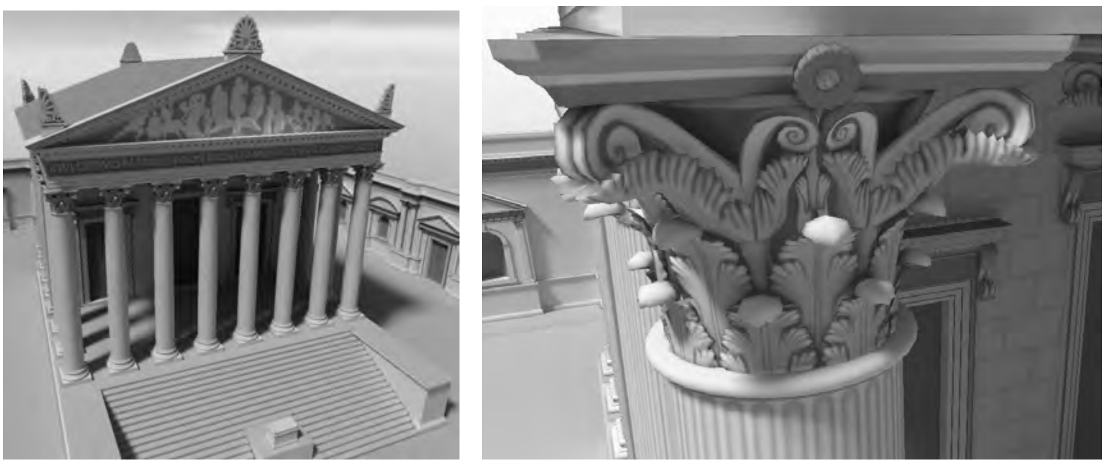


Figure 1: Screenshots from the interactive 3D visualization of Roman Cologne: temple and a detailed view of a capital.

1. INTRODUCTION

Virtual 3D reconstructions of archaeological excavation sites, artefacts, and architecture play an important role preserving our cultural heritage for future generations. Used in combination with interactive visualization technology, they become a powerful tool to support scientific discussions among experts and to present important archaeological facts to broad audiences using museum and edutainment applications. However, the creation and interactive exploration of high-detailed 3D reconstructions is still a challenge for the community (Santos et al. 2006, Kuchar et al. 2007) and most projects finally result in still images, pre-rendered animation videos, or QuickTime® panoramas as a compromise between visual quality and interactivity (e.g., Debevec, P. E., 2005, Almagro et al. 2006).

As an interdisciplinary team, consisting of archaeologists, designers, and computer graphic engineers, the project „*Visualization of Roman Cologne*'' started, with the vision to overcome these limitations. Our aim was to construct high-detail virtual 3D models and a visualization framework that enables their interactive exploration and presentation (Figure 1). This paper presents the results of the combined expertise of all teams. It can be read as a guideline for similar future projects, e.g., to setup a collaborative content creation process, select appropriate data exchange formats, or to apply the presented visualization and optimization techniques to other domains of virtual archaeology.

The paper is structured as follows: Section 2 describes the complete creation process from archaeological reconstructions to the 3D visualizations. Section 3 discusses the optimizations of geometric complex reconstructed building models to increase their applicability for an interactive visualization. Section 4 presents the flexibility of our framework that is completely configurable via XML descriptions and can be coupled with Adobe® Flash® widgets to present secondary information using multimedia content. Section 5 presents enhanced visualization techniques that can be used to ease the communication of specific archaeological aspects to the user. Finally, Section 6 summarizes the results and gives an outlook to future work.

2. CONTENT CREATION PIPELINE

To start the collaborative work, we first setup a content creation pipeline, define the data exchange formats, possible use-cases, and the roles for each team. This process should guarantee that all experts within the teams can operate in their domains without restrictions. Additionally, the pipeline is designed to facilitate three important aspects:

1. **Content preservation**: The representation and encoding formats of the accumulated primary (3D models) and secondary data (multimedia content) has to ensure its accessibility for future usage.
2. **Extensibility**: The established pipeline should support the integration of upcoming ideas and new technology in all its stages during the project.
3. **Re-usability**: The developed framework should be robust, flexible, and easily adaptable to enable its application in other interdisciplinary visualization projects.

Figure 2 illustrates the workflow that is described in detail as follows.

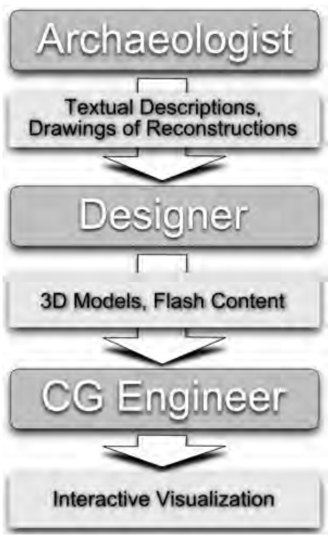


Figure 2: Illustration of the simplified content creation process with the involved teams and the data flow.

2.1 Scientific Reconstruction

To create a complete 3D model representing the ancient town of Cologne, a number of single roman structures and building elements have to be reconstructed first. The reconstruction of these elements, their combination, and arrangement was done by the archaeology experts. For this task, known facts from previous research, results of actual publications (e.g., Irmler 2004), as well as recent finds of the local department of antiquities of the Romano-Germanic Museum Cologne were considered. Only well-studied buildings, with a scientifically evaluated shape or floor plan, were reconstructed in high detail. Thereby, analogies to reconstructions, done before, were used to derive new reliable 3D models. For all other elements only simple shapes were selected for the reconstructions to communicate the missing evidence.

Additionally, a digital terrain model (DTM) (Weibel and Heller 1991) of the ancient Cologne with building sites and streets was reconstructed to embed the 3D buildings later on. It was derived from a DTM of the present Cologne, whose geo-reference was adapted to the location of finds as well as to known morphological changes in the past. Furthermore, the ancient settlement area was bounded by a plateau and a city wall; structures that can be partially recognized in today's cityscape (Figure 3).

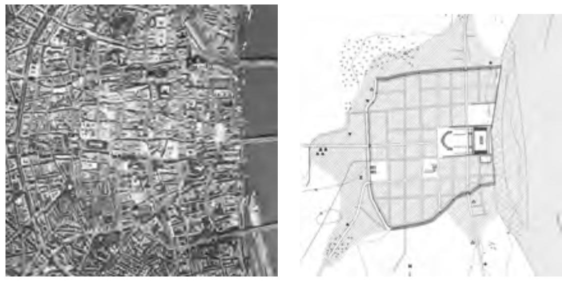


Figure 3: Aerial photographs of Cologne today (left) are used as one additional source for the reconstruction of the digital terrain model of the ancient town (right).

The reconstruction results represent the scientific fundament and are used by the design team to create the virtual 3D models. Therefore, archaeologists deliver all suitable data such as textual descriptions, photographs of artefacts, and highly detailed 2D computer aided design (CAD) drawings (in DWG file format) of building parts and their arrangement to the designers.

2.2 3D Model Creation

Based on this scientific evaluated material, designers create virtual 3D building models for the real-time visualization and Flash® content for the presentation of related information. Thereby, one challenge for designers is to balance the classical trade-off between visual quality and suitability for real-time rendering to meet the requirements of the two other groups. On one hand, archaeologists demand for maximum visual quality for every detail. On the other hand, computer graphic engineers rely on low polygonal representations for fast rendering on graphics hardware, to ensure interactivity for the expected number of high-detail building models.

The 2D CAD reconstruction drawings, imported into a 3D modeling tool, serve as blueprints to model single building pieces. Because buildings in Roman architecture can contain a large number of small ornamental elements, which would result in a huge number of polygons within the 3D reconstruction, designers have to work anticipatorily. To reduce the geometric complexity, rounded structures, represented by spline curves (Foley et al. 1990) in the CAD drawings, are approximated by polygonal counterparts. Generally, all individual objects are constructed with the least possible amount of polygons while remaining the original shape.

After a library of 3D elements (e.g., capitals, columns, doors) was created, these structures are combined to complete buildings and provided with 2D texture-maps representing the respective materials. These material textures are based on color templates prepared by the archaeologists. Further, static lighting conditions are assumed to enhance the 3D impression of the models and to accelerate rendering. Therefore, the material textures are combined with light maps, which are derived from global illumination simulations. Finally, all building positions are geo-referenced and scaled with respect to the overall 3D scene.

To exchange 3D models a data format has to be selected that meet the following three criteria:

1. It has to be supported by digital content creation (DCC) tools used by the designers (Cinema4D, 3DSMax).
2. To provide content preservation, the format should be extensible, future proof, and store information lossless.
3. The format should support a minimum set of standard features (e.g., geometric transformations, color and material definitions, geometry instancing, and external references for reusing building elements.

To enable the use of state-of-the-art rendering techniques supported by modern graphics hardware, e.g., multi-texturing or shader programs (Kessenich 2006), the support of related features by the format is additionally desirable.

We decided on using the *Collada* exchange format (Arnaud and Barnes 2006, Khronos Group 2008). It fits all of above requirements, is based on an XML scheme, has an open specification, and is supported by a large number of major 3D hardware and software companies.

2.3 Interactive 3D Visualization

Compared to static 2D illustrations, interactive 3D visualizations, presenting the Roman buildings within their context, enable archaeologists to discuss or validate their hypotheses with a direct access to the spatial situation. Thus, it become possible to analyze the arrangement of buildings related to the terrain, their mutual visibility from a pedestrian perspective, or the connectivity regarding to the path network. To support these tasks in virtual archeology applications, two conditions have to be considered:

1. Scientific users demand for free navigation, allowing them to inspect every detail in the complete 3D scene. They do not accept restrictions to particular areas or viewing angles, such as used for optimizations in 3D computer games.
2. The models delivered to the 3D CG engineers are created with standard DCC tools. Again, compared to specialized level editors or content creation tools, this allows only a limited set of optimizations during the modeling stage.

For these reasons, and because Collada is more a flexible exchange format than an efficient storage or rendering format, the 3D models have to be converted in a first step. Therefore, we developed a configurable converter tool chain that applies a set of optimizations to each model to prepare it for efficient real-time rendering. To minimize the loading times this converter uses a simple binary format for persistent storage of the optimized 3D model representations. Nevertheless, the Collada representations are not affected and stay preserved as input data for other visualization applications in the future.

Besides model optimization, the second task of the CG team is to develop an application framework for the interactive presentation, analysis, and exploration of the models. To support the need of different end-user groups, this team is responsible for permanent framework extension with enhanced visualization and adaptive navigation techniques.

3. OPTIMIZATION FOR REAL-TIME VISUALIZATION

3.1 3D Model Optimization

All 3D models pass a number of optimization steps to maximize their real-time rendering performance (Kuehne et al. 2005):

1. A *polygon cleanup* operator removes all obsolete data, such as unused texture coordinates, normal data, or vertex colors. Further, redundant information such as vertex duplicates or degenerated triangles are removed.
2. Next, elements of the scene graph are reordered to *minimize state switches* for the graphic hardware. Afterwards, polygonal representations that using the same material or texture form a group whose elements are rendered together in a sequence.
3. The third optimization adds an *index structure* to the polygonal representations if a large number of triangles sharing the same vertices coordinates. This results in a more compact representation and reduces the allocation of the limited graphic board memory. Afterwards, indices are reordered to *optimize the cache hits* during the vertex processing in the graphic processing unit.
4. In the fourth step, geometries for different objects are merged to batches with a maximum of 2^{16} vertices to improve the rendering throughput.
5. At least, hardware *texture compression* is applied for each texture to reduce the GPU memory consumption and accelerate their transfer from system memory to the graphics board.

3.2 Rendering Optimizations

To improve the rendering performance for a scene constructed out of a number of individual 3D models, a further set of optimizations is applied. Because, illumination was pre-calculated for static lighting and stored combined with material colours in the surface textures, lighting calculations are turned off during interactive rendering. Appropriate shader programs are used to minimize the processed set of vertex and fragment operations. These programs calculate only the model-view and projection transformations for each vertex, and use a single texture look-up to determine the visible colour and intensity for each drawn pixel.

To reduce the triangle count processed by the GPU per frame, we apply standard culling techniques (Akenine-Möller and Haines 2002) and introduce a simple but effective level-of-detail (LOD) mechanism. Thereby, highly-detailed parts of building geometry, e.g., capitals, are omitted during phases of intense user navigation. If the interaction with the scene stops, these elements are rendered subsequently to achieve a depiction containing all details.

4. VISUALIZATION FRAMEWORK

4.1 XML Configuration

Both, the tool chain for model optimization and the visualization application can be completely configured using

XML files. This data-driven approach gives the flexibility to easily create and validate new optimization schemes or 3D visualizations scenarios without programming skills.

The XML file for the 3D visualization enables the definition of different visualization scenarios. For each scenario, multiple camera positions and orientations, the 3D models that should be displayed, and a set of navigation constraints can be defined. The camera settings describe predefined views to important elements of the scene. An animated camera movement can be derived by a consecutively interpolation between two camera settings. Navigation constraints help to avoid getting-lost situations. Therefore, a maximum distance to a point in the scene or a minimum height for the observer position, e.g., above the digital terrain, can be specified.

4.2 Communication with Flash Widgets

The 3D visualization can be controlled by other widgets which show additional multimedia information using Adobe® Flash®. This functionality was developed for the use in museum and edutainment applications. The communication between 3D visualization and widgets is based on a client-server architecture. Decoupling these widgets from the 3D visualization enables the independent creation and modification of additional content without side effects for our framework.

Figure 4: Screenshot of a flash widget that is used to control a 3D visualization screen for the presentation in a museum.

Figure 4 shows a prototype for a museum presentation, where Flash® content is presented on a separate computer with a touch screen interface. In addition to all presentation features provided by Flash®, the user can trigger commands from this screen to load a particular scene, or take the view of a predefined observer position in the 3D view. These commands are communicated using standardized network mechanisms. After such a command call, the 3D visualization informs the Flash® widget about the execution status of the triggered command.

5. ENHANCED VISUALIZATION TECHNIQUES

5.1 Cut-Away Views

Interactive cut-away views are an important visualization technique that reveals the interior of complex models by clipping either occluding parts or outer layers (Li et.al 2007). Usually, these depictions are static and where created by hand, thus the view point and the displayed cuts are fixed. Interactive cut-away views overcome these drawbacks by enabling the user to choose desired cut planes and volumes, while navigating in the virtual environment simultaneously.

To enable this functionality in our framework, we integrate a technique described in (Trapp and Döllner 2008). It supports clipping against a number of arbitrary solid volumes within a single rendering pass. The volume meshes can be designed using standard modelling tools and are positioned into the scene using the XML configuration file. At runtime, rasterized representations of these volumes are created in a pre-processing step. During rendering, every fragment is tested using a volumetric parity test with respect to these representations. The fragment is discarded if its associated position lies inside a specific volume.

Figure 5 (left) shows an example for applying two different cuts to a building model in real-time. It reveals parts of the building footprint and internal structures, such as walls and doorways, which otherwise would be hidden to the viewer. Figure 5 (right) shows a close up on a clipped column with applied capping to convey an impression of solid material on the clip surface.

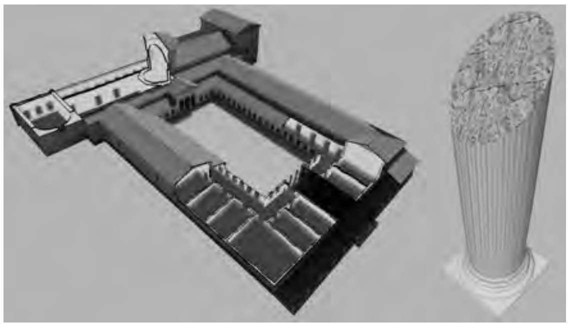


Figure 5: Left: Generated cut-away view of the Praetorium building model using two cut volumes. Right: Single column that is clipped and capped.

The quality of the 3D models, in terms of modelled interior, solid walls, and consistency of polygon orientation, is important for the resulting visual impression. Besides additional configuration issues, the usage of cut-away views introduce a number of challenges and technical implications to our visualization framework:

- **Shading Discontinuities:** Figure 5 (left) shows possible discontinuities in shading for areas inside and outside the building and shadows. This is caused by the pre-lighted object textures that are only valid for views from outside. This effect can be compensated partially by using image-based global lighting approaches that approximate ambient occlusion (Luft et.al 2006).
- **Performance Decrease:** The applied clipping technique is fill-limited, i.e. the runtime performance depends on the number of fragments tested against the volumetric raster representation. To avoid a significant performance decrease, this technique is currently applied to parts of the full scene.
- **Advanced Occlusion Culling:** To deal with the rendering of hidden but potential visible geometry, the framework has to apply more advanced occlusion culling mechanisms (Mattausch et al. 2008).

5.2 Automated Label Placement

Labels are used to communicate short textual information that can not be derived from the depiction of the 3D model. Typically, they identify object parts, show measurements, or give hints to secondary material, e.g., using website URLs. For single views and a small amount, labels are can be placed by hand. However, interactive environments call for an automated process that calculates adequate label positions in real-time. Thereby, each label should be placed readable and in a way that allow users to assign it with the related object unambiguously. Further, a label should neither overlap other labels nor important elements of the depiction and should not be occluded by elements closer to the observer.

We use the approach presented in (Maass and Döllner 2008) that seamlessly integrate labels into the 3D scene instead of placing them in a separate layer, where they superimpose the depiction (Figure 6). Therefore, hull bodies that generalize the complex object geometry are used to define all areas suitable for embedding a label. The simplest hull variant covers all large areas of a building with a set of rectangles. Sample points that serve as label candidate positions are equally distributed over these rectangles. To allow a fast determination which position at which rectangle provide the best quality for a given viewing direction, the visibility of hull bodies, hull elements, and sample points is tested hierarchically. For each rectangle a score for the orientation and visible area is calculated to sort them with respect to their quality potential. Afterwards, all rectangles are iterated using this order until a position that allows a visible label embedding is found.

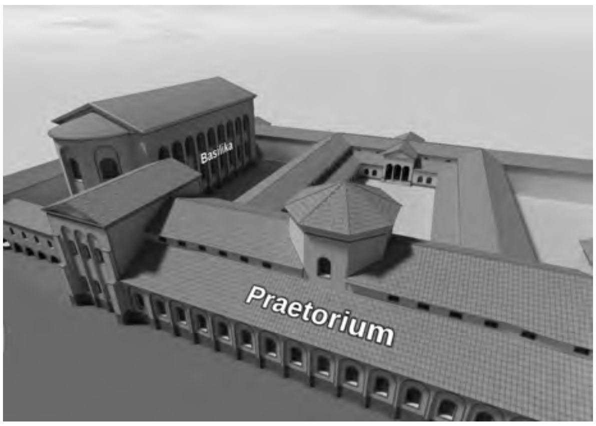


Figure 6: Embedded annotations dynamically placed at large areas that are oriented towards the viewer are used to communicate textual information about the building models.

5.3 Non-photorealistic Rendering

The style used to create an image has a huge effect on how the depiction is interpreted. Photorealistic renderings tempt people to believe they look at photographs of real world objects. For virtual archaeology this kind of presentation is often not sufficient. Archaeologists require the ability to express uncertainty such as missing evidence or accuracy in the reconstructions. Artist often use sketchiness in their illustrations to communicate such uncertainty. In computer graphics the generation of stylistic drawings belongs to the area of non-photorealistic rendering (NPR). A large number of different NPR techniques have been published in the research literature. An overview can be found in the textbooks by (Gooch and Gooch 2001) and (Strothotte and Schlechtweg 2002).

In (Roussou et al. 2003) watercolor and pen-and-ink representations are used to communicate uncertainty in archaeological models. For our purpose, it is important that the NPR technique can be applied in real-time and that the integration into our framework can be done without major structural changes. Therefore, we choose an image based abstraction technique that can be applied as a post-processing effect to the intermediate photorealistic rendering result (Kyprianidis and Döllner 2008). The technique smoothes low-contrast regions while preserving edges and enhance salient important edges (Figure 7).

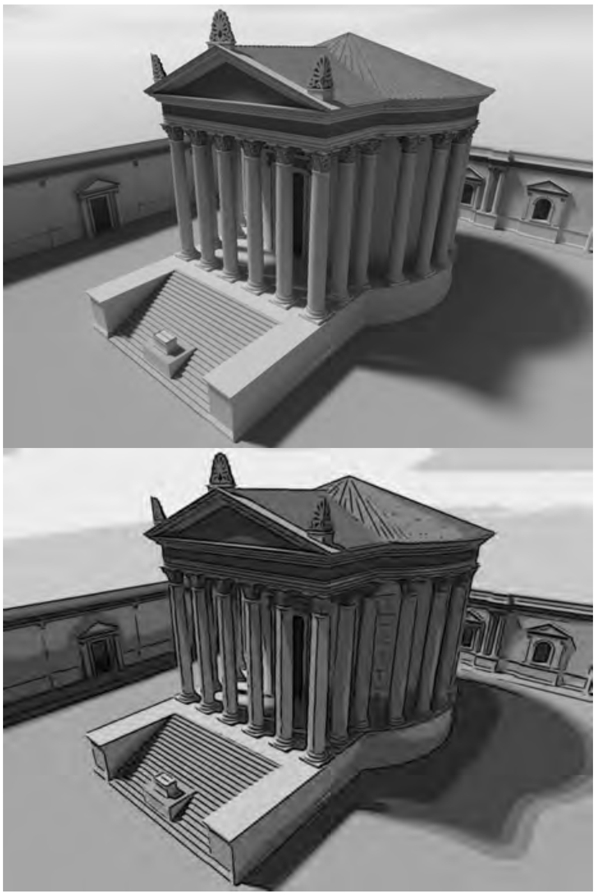


Figure 7: 3D model of a temple. Top: Photorealistic rendering. Bottom: NPR rendering to communicate missing evidence.

6. RESULTS AND OUTLOOK

The established content creation process as well as the chosen data formats proven themselves in practice. The developed framework allows the interactive exploration of the virtual reconstructed Roman Cologne. Figure 10 shows a screenshot from the interactive 3D visualisation containing all currently reconstructed buildings and the terrain model. For most of them, modelling, texturing, and lighting simulation is finished. A complete list with the geometric complexity and the amount of RGB texels used for each model is given in Table 8. Figures 11-13 show additional screenshots from the real-time visualization to give an impression about the high visual quality of the 3D building models.

Model	Complexity		
	#Vertices	**#Triangles**	**#Texels**
Ara Ubiorum	1,652,996	2,197,182	49,670,283
Cryptoporticus	1,477,628	1,890,087	49,670,283
Dionysos Villa	143,296	236,276	33,113,522
Forum Basilica	438,132	521,718	16,556,761
Insulae	350,298	522,662	62,694,105
Praetorium	416,822	600,494	74,615,652
Round Temple	379,655	540,655	49,670,283
Temple	2,378,845	3,210,162	66,227,044
Market Halls	237,648	269,912	16,556,761
Terrain	52,811	92,223	33,113,522
City Wall	248,672	483,174	0
Sum	7,776,803	10,564,545	451,888,216

Table 8: Geometric and visual complexity of the 3D models used in the interactive visualization.

The applied optimizations enable our framework to render the complete model (without omitting high-detailed building parts) at interactive frame rates. Table 9 shows our measurements on four different test systems.

Configuration	Screen Resolution	FPS
Athlon 64 X2 Dual-Core 4200+ 2.21 GHz, 2GB RAM, GeForce 8800 GTS 640 MB	1600x1200	20
Intel Xeon E5345 2.33 GHz, 3GB RAM, GeForce 8800 GTX 512 MB	1600x1200	20
Intel Core 2 Duo X6800, 2.93 GHz, 3 GB RAM, GeForce 7950 GT 512 MB	1600x1200	12
Intel Core 2 Duo T7700 2.4 GHz, 3GB RAM, NVidia Quadro FX570M 128 MB	1400x1050	8

Table 9: Average rendering frame rate per second (FPS) measured on three different test systems for the whole scene of Roman Cologne.

As next step, we are going to complete the city of Roman Cologne. For the future, it is planned to further refine and extend the city model. One concrete idea is the reconstruction of the street of tombs, which was located outside the gates of the city walls. However, with increasing model complexity, we have to switch to out-of-core rendering techniques (Dietrich et al. 2007), because it is currently near to the limits of today's graphics hardware. Furthermore, we plan extend our real-time visualization with high dynamic range light simulation (Kuchar et al. 2007) to allow an adapted presentation of interior building structures.

Additionally, we want to make the huge amount of existing scientific material accessible within the project. Therefore, we plan to extent the framework to interlink elements of the buildings with material from existing scientific data bases, e.g. the *Arachne* database (www.arachne.uni-koeln.de).

7. REFERENCES

Almagro, A., Rodríguez, C., González, M., Zúñiga, I., 2006. The Alcazar of Seville in the 14th Century, *7th Symposium on Virtual Reality, Archaeology and Cultural Heritatge (VAST)*, pp. 15-22.

Akenine-Möller, T., Haines, E., 2002. *Real-Time Rendering*, 2. ed.. A K Peters Ltd., Wellesley, Mass..

Arnaud, R., Barnes, M. C., 2006. *Collada - Sailing the Gulf of 3D Digital Content Creation*. A.K. Peters Ltd., Wesley Mass..

Debevec, P. E., 2005. Making "the parthenon", *Invited paper of VAST 2005*, Pisa, Italy. http://www.debevec.org/Parthenon/debevec-vast2005-parthenon.pdf.

Dietrich, A., Gobetti, E., Yoon, S.-E., 2007. Massive-Model Rendering Techniques., *IEEE Computer Graphics and Applications*, 27(6), November/December 2007, pp. 20-34.

Foley, J. D., van Dam, A., Feiner, S. K., Hughes, J. F., 1990. *Computer Graphics: Principles and Practice*, 2nd ed., Addison-Wesley Professional.

Gooch, B., Gooch, A. A., 2001. *Non-Photorealistic Rendering*, AK-Peters.

Irmler, B., 2004. Rekonstruktion einer Porticusordnung an der Rheinseite in Köln. 77–102, *Kölner Jahrbuch 37*, Gebr. Mann Verlag, Berlin.

Kessenich, J., 2006. *The OpenGL Shading Language*, Version: 1.20 Document Revision: 8, Manual, September 2006.

Khronos Group, 2008. *Collada 3D Asset Exchange Format* http://www.khronos.org/collada/, (accessed 20 Aug. 2008).

Kuchar R., Schairer T. and Straßer W., 2007. Photorealistic Real-Time Visualization of Cultural Heritage: A Case Study of Friedrichsburg Castle in Germany, in: *Eurographics Cultural Heritage Papers*, pp. 9-16.

Kuehne, B., True, T., Commike, A., and Shreiner, D., 2005. Performance OpenGL: Platform Independent Techniques. In *ACM SIGGRAPH 2005 Courses* (Los Angeles, California, July 31 - August 04, 2005). J. Fujii, Ed. SIGGRAPH '05. ACM, New York, NY, 1.

Kyprianidis, J. E., Döllner, J., 2008. Image Abstraction by Structure Adaptive Filtering, *Proc. EG UK Theory and Practice of Computer Graphics*, pp. 51–58.

Li, W., Ritter, L., Agrawala, M., Curless, B., Salesin, D, 2007. Interactive Cutaway Illustrations of Complex 3D Models, *Proc.s of the ACM SIGGRAPH Conference* 2007, pp 31.

Luft, T., Colditz, C., Deussen, O., 2006. Image enhancement by unsharp masking the depth buffer, *Proc. of the ACM SIGGRAPH Conference* 2006, pp. 1206-1213.

Santos, L. P., Coelho, V., Bernardes, P., Proença, A., 2006. High Fidelity Walkthroughs in Archaeology Sites, *6th Int. Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, Pisa, Italy, pp. 453-458.

Maass, S., Döllner, J., 2008. Seamless Integration of Labels into Interactive Virtual 3D Environments Using Parameterized

Hulls, *4th International Symposium on Computational Aesthetics in Graphics*, Lisbon, pp. 33-40.

Mattausch, O., Bittner, J., Wimmer, M., 2008. *C*HC++: Coherent Hierarchical Culling Revisited, *Computer Graphics Forum*, 27(2), pp. 221-230.

Rost, R. J., 2006. *OpenGL Shading Language*, 2. ed.. Addison-Wesley, London.

Roussou, M., George Drettakis, G., 2003. Photorealism and Non-Photorealism in Virtual Heritage Representation. , *Proc. of the 4th Int. Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, Eurographics, pp. 51-60.

Strotthotte, T., Schlechtweg, S., 2002. *Non-Photorealistic Computer Graphics: Modeling, Rendering, and Animation*, Morgan Kaufmann.

Trapp, M., Döllner J., 2008. Real-Time Volumetric Tests Using Layered Depth Images*, Proc. of Eurographics* , pp. 235-238.

Weibel, R. Heller, M., 1991. Digital terrain modelling, In D.J. Maguire, editor, *Geographical Information Systems: Principles and Applications*, Longman Scientific & Technical, pp. 17–90.

8. ACKNOWLEDGEMENTS

This work has been funded by the German Federal Ministry of Education and Research (BMBF) as part of the InnoProfile research group '3D Geoinformation' (www.3dgi.de).

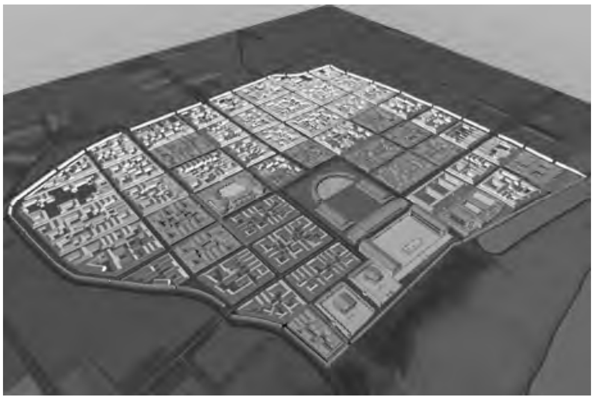


Figure 10: Overview of the 3D visualization containing all currently reconstructed buildings of Roman Cologne.

Figure 11: Screenshot from the real-time visualization with a view from within the Praetorium model.

Figure 12: Screenshot from the real-time visualization showing the details of the Ara Ubiorum model.

Figure 13: Screenshot from the real-time visualization showing the details of the Cryptoporticus model.

SKETCHING TECHNIQUES IN VIRTUAL REALITY: EVALUATION OF TEXTURING STYLES IN AN URBAN PLANNING MODEL

B. Stahre [a, *], S. van Raalte [b], I. Heldal [c]

[a] *Dept. of Architecture; Chalmers University of Technology; Gothenburg; Sweden - bea@chalmers.se
[b] Vianova Systems Sweden AB; Gothenburg, Sweden - susanne@novapoint.se
[c] Dept. of Technology & Society; Chalmers University of Technology; Gothenburg; Sweden - ilona.heldal@chalmers.se

KEYWORDS: Virtual Reality, Textures, Sketch-styles, Presentation, Infrastructure, Design Process, 3D, Colour Appearance

ABSTRACT:

Today's design processes involve many participants and demand new methods of communication for designers, developers and users. Using Virtual Reality (VR) as a design tool has become increasingly common. Regarding visual expression, photorealism is often strived for, even though this might not always be in accordance with the main purpose of the visualization-project. The research behind this paper regards the development of VR as an architectural design tool that can be used throughout the entire design process, from sketch to final presentation. Here we present results from a comparison between different sketching styles and photorealism in the texturing of an urban planning model. The aim is to show how sketch-like expressions in the texturing of a model can clarify and simplify the understanding of a building project (area plan, housing area, road design). The target group is located within municipalities (architects, planners) and the consultant industry (road designers/engineers, landscape architects). Data was collected from questionnaires answered by 20 participants, all of them professional users of VR. They assessed the experience of the texturing styles in the test model on desktop-PCs. The results revealed important differences and similarities in the perception of the sketching styles vs. the photorealistic style. The evaluation revealed a desire for more sketch-like expressions supporting conceptual design thinking. Even so, models should provide a high level of detail and good spatial experience. Aesthetic factors are considered important. The results contribute to a better understanding of technical and aesthetic limitations of photo-realism in VR.

1. INTRODUCTION

Using Virtual Reality (VR) as a tool for visualization in the design process is becoming increasingly common, and the advantages of using this new medium for communicating ideas are many. However, computer generated visualizations can in some ways also increase the distance between the designer and the representation of the design ideas (Brown, 2003). Visualizations in the early stages of the design process which aim to look as finished and realistic as possible may limit creative thinking, since the flexibility for the final representation is settled too early (Oh, 2006). Using photorealism as a standard representational style in visualizations raises issues concerning the visual expression. In today's VR-visualizations photorealism appears to be the standard representational style, irrespective of what the visualization is intended to show. It can also be discussed if a photorealistic expression in the visualizations can be misleading regarding how finished the depicted design plans are, when on the other hand, the photorealistic level in interactive architectural models is often unsatisfactorily low.

For the design, communication and criticism of architecture, architects depend on representations (Bermudez, 1995). Traditionally, sketches have been used by architects to form and convey their ideas throughout the design process. The value of sketches in the development of a design project lies in the intent to illustrate an idea or a concept that should support creative thinking and conceptual design processes. What is important in a sketch is to show the focus of the idea, i.e. to illustrate the different levels of relevance in accordance with the concept. Another advantage of using sketching is that this technique can pedagogically express how developed the ideas are, i.e. how far in the design process the project plans have come. By using sketching the architect has the ability to visualize information in the project that is not yet fully decided or thought through, as well as to visually enhance aspects of the design which it is important to focus on. For VR to be a useful design tool throughout the whole design process, these values need to be considered. The design process is a wide concept with many interpretations. We consider it to be roughly divided into the early design phases of pre-conception (e.g. associative imagery) and sketching/design, and the later phases of presentation and management.

Our research deals with developing VR as an architectural design tool through visual expression, throughout the entire design process. This paper is based on the research project *Sketching Techniques in Virtual Environments (STIVE)* The problems addressed are connected to 1) how a VR-visualization can convey the different levels of information in an idea under development, throughout the whole design process, and 2) how the creative characteristics of an architectural sketch can be conveyed into a VR-visualization and adjusted to the specific pre-requisites of this technique. The aim with this paper is to evaluate how different alternative sketching styles in the texturing of an interactive virtual model can clarify and simplify the understanding of a building project (area plan, housing area, road design) during different stages in the

* Corresponding author.

planning- and design process, from initial sketches to final presentation. The motive for focusing on textures in the visualizations is to enable users to get better control of the visual expression. The questions we will discuss in this paper concern both visualization techniques and the experience of a visualization. By analyzing the results from the study we wish to answer the following questions: 1) Which advantages and problems do the participants relate to VR? 2) Is there a need for different texturing styles to better support the different stages of the design process? 3) To what extent are parameters such as colour, details, and aesthetic values of importance in the interpretation of visualization? These questions will form a part of a more general discussion on representational issues of visualization. In this paper these issues will be discussed from the perspective of architectural research and in relation to current research. Target groups are architects, landscape architects, planners, road designers and others interested in representational issues in VR.

The article is structured as follows: Chapter 2 draws a concise picture of research using VR models and different sketching techniques in architectural planning processes. Chapter 3 presents methodological considerations and the experimental project behind this paper. Chapter 4 includes qualitative and quantitative results from a questionnaire completed by 20 professional users of VR in architecture. Chapter 5 discusses these results from the viewpoint of discovering how different texturing styles and sketching techniques better support design processes. The last section presents conclusions and future work. The term VR-visualization refers in this context to a computer generated architectural 3D-model, which can be used as a base for rendered images, for animations or for an interactive visualization and in which a user can move around in real time.

1.1 Problem Area and Relevant Research

The field of visual representation in VR is broad and includes many disciplines, from imaging and technical areas of expertise to cognitive science and art. Unwin (2007) refers to the literature of architecture when distinguishing three main uses for architectural drawing, that is: as a medium for *communication* (with clients, builders etc), as a medium for *design* (private 'play') and as a medium for *analysis* (to acquire knowledge and understanding). Among architects VR as a tool for visualization is above all used in order to communicate ideas (Setareh et al. 2005). In this paper we will mainly focus on visualizations as a means of *communication*.

As mentioned in the introduction, the value of sketches in the development of a design project lies in the intent to illustrate a design idea or a concept rather than showing the real world setting as it is. However, two dimensional drawings can sometimes be hard for laymen to interpret, and thus are not optimal for architects as a means of communication. For example they have difficulty providing a correct impression of scale as well as perspectives of every space from all angles, which is something that VR-visualizations facilitates (Savioja et al. 2003). Communication to non-specialists seems thus to be made clearer and easier through the use of computer visualizations, i.e. both interactive representations and rendered realism. (Neto, 2003) It is however still difficult to design and implement a trustworthy virtual environment, even with today's progressive technology. The traditional assumption has been that by making interactive models look as visually realistic as possible, more believable virtual experiences have been created (Drettakis et al, 2007). But still, visual realism is hard to obtain, mainly due to the complexity and richness of the real world. Neto (2003) stresses the importance of putting great care and critical information into the creation of visual computer technologies. He states that interactive models used in planning and design practice, due to still being too artificial looking, mostly lack the necessary believability to be accepted as reliable tools for evaluating the proposed urban or architectural space (Neto, 2003).

It appears that the research area of architectural representation in VR needs to be further studied in order for VR models to be used to correctly convey ideas throughout the design process. Kwee (2007) notes that the area of digital architectural presentations focuses on the technology's provision for speed and ease of information retrieval. In the meantime, the quantity and presentation of information in these visualizations are assumed, without proof, to be currently adequate for mediating correct understanding. He states that there still needs to be much rethinking and improvement to consider in order to understand the potential of digital visualization for architectural presentations. (Kwee, 2007) Although exploratory usability-oriented studies involving VR-programs have been carried out (e.g. Panagiotis et al, 2006), very few studies have been reported on the role that VR plays, and could play, in ongoing environmental planning contexts (e.g. Heldal et al. 2005). Balakrishnan et al. (2007) observe that physical objects rather than the spatial experience are emphasized in common digital tools for design visualization. In current rendering technologies great achievements are made in representational similarity through increased photorealism. Accordingly the challenge lies in the experimental concordance with a corresponding real space. Balakrishnan et al. state that more work needs to be done exploring current tools of digital representation, in order to improve aspects related to the experience of a simulation (Balakrishnan et al. 2007).

In relation to the technical development for visual rendering, there are new techniques that support non photorealistic rendering (NPR), e.g. by generating textures with boundary effects (Ritter et al, 2006) and creating 3D shapes from 2D contour sketches (Karpenko et al, 2006). These works use *given* background information (e.g. from topology databases) where the naturalism of certain features (e.g. boundaries, volumes – provided by the databases) are important. Thus the form of final products can be predicted. Contrary to this, our project aims to find the necessary information (e.g. simple lines and shapes) that supports creativity to obtain new solutions. Here the new forms should support conceptual design thinking and not be limited by predefined structures. Bermudez (1995) encourages us not to concentrate our investigations on the computer's power to do what we already know how to do, but instead to focus on the distinctive features and uniqueness of digital media. He states that the unique ways in which electronic representations address architectural issues, elements, ideas and design problems need to be dealt with. (Bermudez, 1995)

In the process of forming a design idea there is a need for different expressions and levels of representation. The visualizations also have to interpret some issues exactly (e.g. Drettakis et al, 2007), while just sketching others (Lange, 2005). The precision of representation can differ from one user group to another or from one design phase to another (Al-Kodmany, 2002). According to Brown (2003), in different phases of the design process different types of representations appear to serve well, and one challenge lies in how best to integrate the different computer generated representations into the design process. Too much detail and visual realism in representations at the initial stages of the design process is often not necessary and can even be misguiding, since that information will not be decided on until later on in the process (Neto, 2003). Several experiments have been conducted on comparisons with the human vision response to computer generated architectural images with a conventional expression of hard edges and straight lines to a hand drawn expression, with wobbly lines (Van Bakergem and Obata, 1991; Brown and Nahab, 1996; Bassanino, 1999). Results from these studies showed that the images with a hand drawn expression, depicting the same building and containing the same data as the hard edged images, were rated higher on qualitative factors such as stimulation and interest (Brown, 2003). Brown notes that "Such images, whether produced manually or via computer hardware and software appear to have more worth attached to them and are regarded more highly as stimulating and pleasing architectural objects." (Brown, 2003)

2. EXPERIMENTAL PROCEDURE

The research project STIVE was conducted between 2007-08 and is a collaboration between the Dept. of Architecture at Chalmers University of Technology and the software developer Vianova Systems Sweden AB. The target group for the study was located within the municipality (architects, planners) as well as within the consultant industry (engineers, landscape architects). The research approach is primarily design-based (Groat and Wang, 2002; Billger and Dyrssen, 2005) focusing on elaborations with texturing in an interactive visualization. A selection of textures were applied in an environmental interactive test model, which was designed to present the different texturing styles from various distances and in varying environmental contexts (housing area, roads and nature). Together with a questionnaire and a user manual, the test model was distributed to the selected group of participants for evaluation. The results were then compiled and analyzed.

2.1 Set-Up

The test model was designed in the infrastructure design software Novapoint Virtual Map[**], which is an add-on modeling and visualisation application on the AutoCAD platform. The choice of software was connected to the need for the participants to be able to switch between the texture styles in a simple way, though they were restricted from changing the actual structure of the model. The participants assessed the task on desktop PCs. As part of the set-up, the participants were required to download the model from the Vianova ftp site and install the accompanying Style-library, which contained the seven different texture styles. Before beginning with the actual evaluation, they were asked to get acquainted with the model for a few minutes. When evaluating the model the participants were asked to go through different viewpoints in each texture style, before answering the questions in the accompanying questionnaire.

In the study, 20 participants took part, both architects and civil engineers, all of them professional users of VR in environmental and architectural contexts. Since the participants were located all over Sweden, the questionnaire and accompanying information had to be sent to them by mail and the other correspondence was handled via e-mail and phone. In order to ease the set-up and the viewing of the model for the participants, one important condition was that they would already be familiar with Novapoint Virtual Map and have access to this software.

2.2 The Demonstration Model

The study was based on a VR model containing built environment, infrastructure and landscape (see Figure 1). A selection of textures in different artistic styles was applied in the model and shown to the participants. The virtual setting which the model displayed consisted of different environmental contexts in a typically Swedish landscape in summer time. Since the context was to be usable for both landscape architects, architects and planners the virtual setting included parts relevant for each profession. One part of the model showed a housing estate, consisting of both single houses and blocks of flats of different sizes, detail and expression. Another part described different types of roads, while the third part described countryside displaying various trees, bushes, flowers and ground materials. In order to simplify the task for the participants, a selection of different view points, showing what was relevant in the model, had been preset. One important criterion was that the model should contain common types of objects that a VR-visualization consists of. These objects include *billboards* (i.e. trees, bushes, people), *buildings* (i.e. facades of a selection of different building types), *ground material* (i.e. gardens, meadows, fields, farmland, woodland etc), *roads* (i.e. paving, slopes, ditches) and *back-drops* (i.e. the edge of a forest, wood fences etc). Another basic criterion was that the model should be equally well represented in all of the included texture styles.

Figure 1: The demonstration model, showing 1) built environment, 2) landscape and 3) infrastructure

2.3 The Textures

Different artistic texture styles were developed and applied in the model. The starting point for the elaborations was the original photorealistic textures in the model. Important in the elaborations was that the textures should work in different stages of the design process. Variations in detail, expression and abstraction were considered. For example, textures containing only outlines or coloured surfaces were assumed to work better in the beginning of the design process, where the concept rather than the details is most important, while the textures containing more details should

** http://www.novapoint.se/produkter.asp/id/30/LID/12627 be suitable for later stages. The general aesthetic expression of the textures was considered more important than, for example, correct colour reproduction in each texture, and a certain amount of artistic freedom was allowed. The elaborations concerned both colours and greyscale. A large part of the elaborations consisted of finding a suitable balance between textures on vertical and horizontal objects in the model. The final selection of texture styles was to include a variety ranging from less to more detailing, more abstract and artistically free styles to realistic ones, and monochromatic to polychromatic ones.

The final selection of texture styles came to include *Realism, Colour, Greyscale, Contour (Colour)*, *Contour (Black and White), Graphical* and *Sketch* (see Figure 2a and 2b). The styles were defined as the most relevant by two experts in the STIVE-project. The textures of *Colour* and *Greyscale* were poly- respectively monochromatic surfaces and contained the fewest details, followed by *Contour (Colour)* and *Contour (Black and White),* which were surrounded by black borders. *Contour (Colour)* contained slight variations in colouring while *Contour (Black and White)* was purely black and white, imitating pen and paper. We thought these styles suitable for the early stages of the design process. The textures *Graphical* and *Sketch* were more sketch-like and artistic, and contained more detailing. They were assumed to be more suitable for presentational use in later stages of the design process. *Realism,* containing most detailing, was the default photorealistic texture from which we started the elaborations.

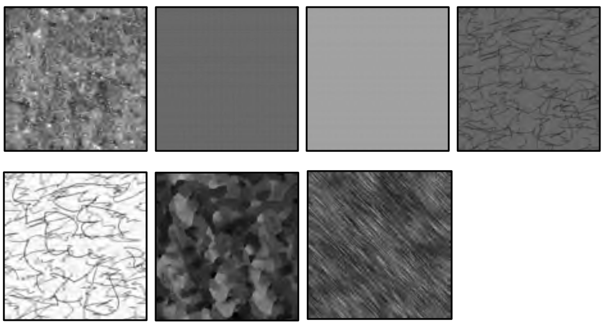


Figure 2a: Ground textures. From top left: *Realism, Colour, Greyscale, Contour (Colour), Contour (Black and White), Graphical* and *Sketch*

Figure 2b: The texture styles on billboards. From top left: *Realism, Colour, Greyscale, Contour (Colour), Contour (Black* and *White), Graphical and Sketch*

Most of the chosen final textures were designed in Adobe Photoshop, apart from the *Graphical* style and the *Sketch* style which were drawn by hand and then manipulated in Photoshop.

2.4 The Questionnaire

The questionnaire was divided into four sections. The first section concerned the professional profile of the participants, while the second focused on their use and experience of working with computers in general. Their experience of working with 3D-visualizations and which software they used was encircled. The third section concerned their experience of VR-visualizations. The participants were asked to give their views and comments on the general advantages and problems of today's VR-visualizations, and how they might be improved. They were also asked to comment on how well their companies' VR-visualizations fulfilled their purpose. In section 2-3 free descriptions were used to a large extent, sometimes as a supplement to encircled answers of a yes/no/don't know-character. The fourth and biggest section of the questionnaire concerned the evaluation of the experience of textures in the accompanying test-model. In this section a few more evaluation techniques were added:

1. *Free description of each style.* The participants were asked to describe, with one or two words, the experience of each texture style.
2. *Motivated semantic differential scaling.* The participants were asked to mark the importance of different characteristics (colour appearance, detailing and aesthetics) for the textures on an open 7-grade scale and to motivate the markings.
3. *Visual evaluation of the model.* The participants were asked to encircle the texture style best suited for each component in the model (ground, billboards, buildings, road, and side-scenes).

The questionnaire was composed of both qualitative and quantitative questions, which were important complements to each other.

3. RESULTS

Here we will present the results from the different parts of the questionnaire.

3.1 Part 1 and 2 of the Questionnaire: Target Group and Computer Usage

Part 1 of the questionnaire concerned the professional profile of the participants, while Part 2 concerned their usage and experience of working with computers. The questionnaire was sent to 34 participants, from among whom 20 answered and returned it. All of the participants were occupied within the municipalities or the consulting industry, with work concerning planning and design of architecture and infrastructure. The group of participants came to consist of 10 architects (5 female and 5 male) and 10 engineers (4 female and 6 male). The age of the participants varied from 28 to 57 years, with an average of 37, 5 years. The participants were all used to working with computers, with an average experience of 15, 5 years and 30-40 working hours per week in front of a computer. They were all familiar with AutoCAD and most of them also with other 3D-visualization software. On the question of how often they worked with 3D, 10% answered "frequently", 60 % answered "sometimes" and 25 % answered "seldom". On the question of who made the visualizations that they used, all of the participants answered that their own companies created them.

20% also used external visualization consultants. It is important to note that all of the participants in the study were working with visualizations in addition to their ordinary work tasks and that the visualizations were to be an aid for them in their work.

3.2 Part 3 of the Questionnaire: Visualization Technique

Part 3 of the questionnaire concerned the participants' opinions on visualization techniques, focusing on the advantages and disadvantages of VR-visualizations. The answers were all qualitative. The main advantage of using VR-visualizations was considered by most (70 %) to be increased understanding of a project. Many of the participants defined this to be the advantage of being able to mediate ideas to clients or the public. 30 % considered the capacity and content of the visualizations to be other advantages. Many also mentioned the ease of orientation and getting a general view of a VR-visualization compared to 2D. Some comments read: "Inexperienced map readers get a clearer picture of the area, heights, layers etc.", "It is easier to mediate an idea to a client by using 3D-views." and "The whole; that it is all there." The main disadvantages of using VR-visualizations were related to the representation of realism (43 %), the technology (31%), the basic data (16%) and the work effort (10%). Many of the participants commented on the representation of realism as a problem. Some of them specified this to regard *the level* of realism in the visualizations. One wrote: "Some users get more annoyed that everything does not look exactly as it does in reality, instead of seeing the purpose of the model." Another one noted that: "Those who see the model often expect it to be more accurate than it actually is." Other participants commented on wrong interpretations of realistic looking visualizations. One remarked: "A too realistic looking visualization can lock the opinion of what an area / a building can become." Another comment read: "It can be a risk that the public believes that it will become exactly as it is shown – which is not always the case." One wrote: "It [the visualization] is not always regarded as a sketch, but is experienced as more fixed." Regarding the technology and the work effort, time-consumption was mentioned by many as one of the biggest problems. One participant wrote: "Too much energy is taken from the "real" design in order to do a good-looking VR-model." Other problems included the size of the visualizations and usability issues, i.e. difficulties in learning new software, and also that visualizations sometimes are too complicated to make. One participant remarked: "[The visualizations are] difficult to run in real time. I wish they worked more like well-made computer games." Difficulties in finding good basic data (orthophotos, terrain data etc) for the visualizations were also mentioned as a problem.

3.3 Part 4 of the Questionnaire: Evaluation of the Texture Styles in the Demonstration Model

The fourth part of the questionnaire contained both main questions with sub questions and groups of questions. Some participants did not answer one or more of the questions, and therefore the numbers of answers which are accounted for in this paper vary between 16 and 20. Important to consider is the qualitative follow up to each quantitative answer. Also important to note is that participants sometimes stated more than one alternative in their answers. For the analysis we have considered this, and chosen to calculate each answer as 1. If the participant stated two styles as answer to a question that demanded only one alternative, we calculated on 0,5 for each stated style. If they stated three styles we calculated on 0,33 for each etc. This procedure has been consistent throughout the analysis. On the question of which texture style they would most likely consider using as a work tool in all stages of the planning and projection-process (see Figure 3a), the participants answered that they would chose *Sketch* (21%), *Realism* (20 %), *Contour (Colour)* (16%), *Colour* (15%), *Graphical* (13%), *Contour (Black and White)* (11%) and *Greyscale* (4%). Comments revealed that the sketch styles were considered good in order to simplify or distinguish objects in a model, as well as for adjusting and varying the model for a specific need or stage. Realism was commented on as being useful in the latter stages of the design process. On the question of which texture style they would most likely consider using as a tool for presentation (see Figure 3b), the participants answered that they would chose *Sketch* (27%), *Contour (Colour)* (24%), *Graphical* (16%), *Realism* (15 %), *Contour (Black and White)* (13%), *Greyscale* (3%) and *Colour* (2%). When analyzing the accompanying comments, it is apparent that the sketch styles were considered more usable as a presentational tool than as a work tool. A mix between some of the sketch styles was asked for by some, as well as the ability to combine sketch style objects in an otherwise realistic looking model to enhance new additions to a plan. One participant remarked "Often it is desirable to be able to decrease the realism in a model. If the realism is too high, smaller details with not so much significance get too much attention."

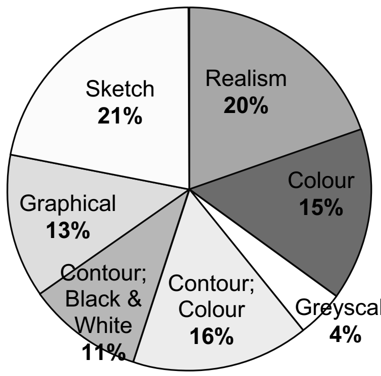


Figure 3a: The texture style best suited as a tool for sketches and projections

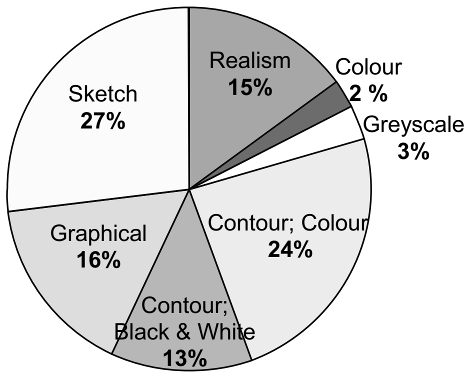


Figure 3b: The texture style best suited as a tool for presentation

The participants were also asked which texture style they considered to work best on each object type (ground, billboards, buildings, road and background) in the model. In their answers *Sketch* was considered to be the best working style for billboards (39,5%), buildings (39,5%) and background (42%).

For roads *Graphical* was most popular (28%) and for ground textures the *Contour (Colour)* (45%) (see Figure 4).

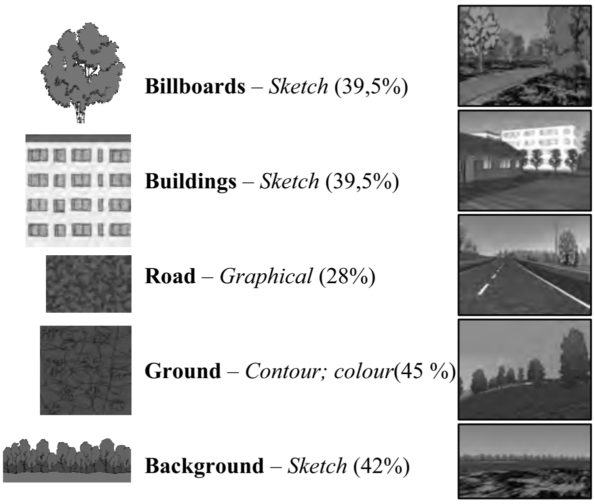


Figure 4: The best working texture style for each object type in the model

The participants were asked to mark the importance that the different characteristics *colour appearance*, *detailing* and *aesthetics* have for the textures on an open 7-grade scale and to motivate the markings (see Table 5).

	Colour appearance	Aesthetics	Detailing
Sum	36	34	19
Mean value	2	1,9	1,1
Modal value	2	2	2
Non-response	2	2	3

Table 5: The importance of colour appearance, aesthetics and detailing for the visual expression of a VR-visualization

The scaling stretched from 3 to -3, where 3 was considered the most affirmative and -3 the most negative. 0 was neutral, but considered in the calculation. The mean value for the importance of *colour appearance* for the experience of the visualization was 2. The modal value (i.e. the highest number of participants marking the same number) was also 2. The mean value for the answers to the importance of *aesthetics* was 1,9. The modal value was 2. The mean value for the importance of *detailing* was 1,1. Here, too, the modal value was 2.

The textures which the participants generally preferred were *Sketch* (41%), *Contour (Colour)* (20%), *Realism* (19%), *Graphical* (11%), *Contour (Black and White)* (7%), *Colour* (2%) and *Greyscale* (0%) (see Figure 6). Evident when analyzing the comments is that the different styles would sometimes benefit from borrowing some of each other's features. One participant remarked for example that *Contour (Colour)* had better textures for buildings and ground, but that the trees and other billboards looked better in the *Sketch* style. *Contour (Colour)* was described as good in being distinctly visible at both long and short distances. *Graphical, Sketch* and *Contour (Colour)* were regarded by some as good in their ability to depict reality in a non-realistic manner. Among the architects 8 out of 10 preferred the sketch styles and 2 preferred realism. Among the engineers 7 preferred the sketch styles and 3 preferred realism.

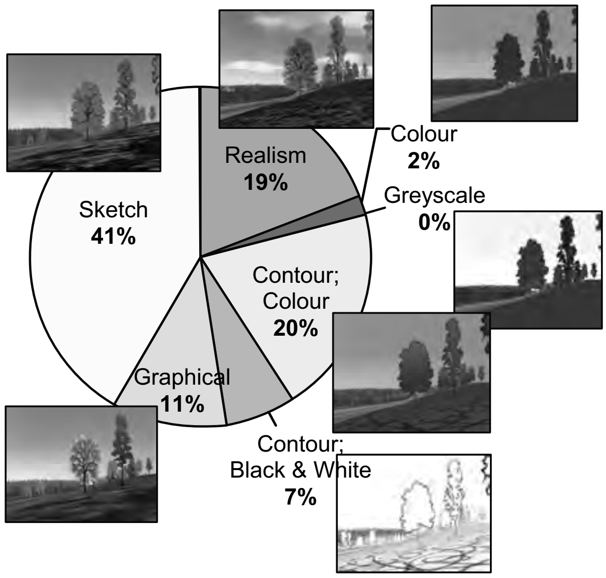


Figure 6: The generally best liked texture styles.

4. DISCUSSION

This discussion concerns the experience and evaluation of the test model in relation to current research on representational issues in interactive visualizations, from the viewpoint of architectural research. The questions we set up to answer were: 1) Which advantages and problems do the participants relate to VR? 2) Is there a need for different texturing styles to better support the different stages of the design process? 3) To what extent are parameters such as colour, details, and aesthetic values of importance in the interpretation of a VR-visualization?

The results gave us distinct answers to the first question. The participants confirmed what earlier studies have already acknowledged as advantages of using VR, such as an increased understanding of a project. Above all, the results point out the potential to mediate ideas to clients / the public. Other benefits of using VR were allowing dynamic content of the visualizations, understanding volumes, and better support for orientation. Regarding the problems of using VR, important results for this study include the representation of realism, which many participants considered a problem; above all the level of realism in the visualizations. The participants also considered it to be difficult to find good basic data (maps, orthophotos, terrain data etc) for the visualizations. A large part of what was considered problematic in using VR was related to technological issues, above all time-consumption and the overly large size of the visualizations. Usability issues and work effort were also pointed out as problems, i.e. difficulties in learning new software and visualizations being too complicated to make. This study also proved that more attention needs to be given to considering appropriate texturing styles in the different stages of the design process. The less complex styles were presumed to fit the early stages of the design process, while the more artistic and detailed ones were presumed to better suit the later stages. The results show that participants disliked certain styles, especially if they were not used to these in their everyday work. For example, the low ratings for the simplest styles (*Colour* and *Greyscale*) revealed that we misjudged these styles as fitting the early phases of the design process. The difference of the gradient, thus a more varied colour range, and the black borders were probably the reasons for the otherwise similar *Contour (Colour)´s* much higher rating. The third stated question concerned the impact that the characteristics *colour*, *aesthetic values* and *detailing* have for the interpretation of visualizations. The high mean and modal values, in combination with the overall disregard for the monochromatic texture styles among the participants, shows the importance of *colour* in VR-visualizations during the whole design process. From the qualitative answers it generally appears to be of lesser significance that the colour appearance is correct, than that the over all impression of the total and combined colour appearance is satisfying and harmonious. Considering the impact that *aesthetics* have, the high mean and modal values indicate that this is very important. Some of the participants commented on the value of presenting a project using a VR-visualization in a non-realistic manner, demonstrating that it is a *proposal* instead of an attempt to visually imitate reality. One participant remarked: "I think a combination of different styles would be useful. Often you want to produce a sketchy expression, so that it does not look too finished. An even more sketchy appearance would be good." This comment may suggest that in some stages of the design process it would be valuable to use sketching tools with more naturalistic information obtained from predefined databases (e.g. Ritter, 2006; Karpenko, 2006). The slightly lower mean value and more scattered modal value of the relevance of *detailing* for a visualization indicates that this was not considered to be of equally high importance. These results are in line with the results from Oh et al (2006). The texture styles with the fewest details, i.e. *Colour*, *Greyscale* followed by *Contour;(Black & White)*, were generally the least popular ones, both as work tools and for presentational use. The exception was a slightly higher rating for *Colour* used as a work tool. This is somewhat surprising since the lack of details in those textures was an attempt to adjust them to fit the first stages in the design phase, where the forming of a concept rather than focusing on undecided details are relevant. One of the participants suggested adjustments for different contexts which would make the styles more applicable on different "zoom-levels".

It is surprising to note that the participants generally considered aesthetic values in a visualization to be of higher importance than detailing. Aesthetic values have always been important in architectural visualizations and many of the participants from an architectural background also expressed a satisfaction with being able to work with aesthetic alternatives. One participant remarked that the sketch styles were good considering that "the architect will recognise himself". The lower ratings of detailing might be connected to Neto's (2003) assumption that too much detailing in the beginning of a design process is more of a hindrance than an aid to the correct perception of a project proposal. From the participant's answers we draw the conclusion that research on representational issues in VR is highly relevant and needs to be further looked into. Users have different purposes with their visualizations and use them in varied contexts, and therefore require different alternatives, which was also stated by Al-Kodmany (2002). Since, among the participants, there was an outspoken need for more sketch-like expressions and the option to choose the levels of detail in the visualizations, the question arises if the striving for photorealism and naturalism is just a current norm irrespective of what the visualization is intended to show. Many participants expressed dissatisfaction with the limited and conformist appearance of today's VR-visualizations. One participant wrote: "What we [in our company] miss in general is the ability to be able to soften the visualization. We liked the sketch style textures and want to see more of that." Another comment read: "These styles are a good step in the right direction. I believe it is beneficial to use a model in different stages and with different expressions." A photorealistic expression in interactive visualizations has sometimes caused clients to interpret the visualized project to be more finished than it actually is, which can thereby lead to unfulfilled expectations and disappointments. On the other hand, in VR-visualizations photorealistic textures in combination with a simplified geometry and lack of realistic interactive light also tends to create an imbalance in the visual expression. This sometimes unsettles users viewing the model, i.e. the model has the ambition to look realistic but does not look realistic enough. Many participants expressed a desire to be able to combine different styles in the same visualization (which was one of our original intentions). Finding a balance between horizontal and vertical textures in the model proved to be difficult, something that some of the participants also commented on. An important remark here is that it is easier to create a trustworthy appearance for textures on vertical objects than for the horizontal areas. This is connected to the lack of borders and the pattern of repetition in the horizontal areas, compared to the frames of single objects for the vertical textures.

In order to see what the two groups of architects and engineers preferred, we did a summary of preferences for the different styles vs. the realistic style. The differences in preference between the two groups were smaller than expected. The group of 10 architects and 10 engineers was however too small for any real conclusions to be drawn. Interesting to note is that the engineers in their qualitative answers were more positive towards the photorealistic texture style, and the architects were more dissatisfied with it. This tendency did not, however, show in any of the quantitative answers.

Considering the approach to this study it is relevant to note that design-based research is not a linear process. Instead it has been necessary to change perspective and go back and forth between theory and practice. The many hours of elaborations can be used to create better design, but above all to increase the understanding of VR's possibilities, as well as to answer questions on when and how we can use this relatively new medium and what the implications of it are for participants with different expertise. The number of participants in this study might be regarded as low. Regarding this, it is important to observe the set-up of the study. For the participants the study was demanding to partake in; both in time and in technological skill. It furthermore required that they would be familiar with and have access to the software as well as fit a specific professional profile. Above all the qualitative answers are very valuable, and through their combination the qualitative and quantitative answers did complement each other well. The variety of questions enables us to identify tendencies for users fitting our profile in their approach and experience of VR-visualizations.

5. CONCLUSIONS AND FUTURE WORK

This paper describes a comparison between different sketching styles and photorealism in the texturing of an interactive environmental planning model. Most current architectural VR-visualizations still struggle with problems related to an aspiration for a photorealistic expression. Models need to be better adjusted to the design phase they describe, and manageable in such a way that the user can modify the visual expression. The evaluation showed a need for greater variety of visual expressions in architectural models, as well as dissatisfaction among many users with the level of photorealism in the visualizations that is possible to obtain today. The results confirm that further research on technological and usability aspects of VR-visualization is needed, e.g. avoiding time-consuming loading times and non-intuitive menus.

For future work there are many improvements and strategies to consider. With the results from the evaluation, we now have a base for further elaborations. More texture styles need to be developed and different levels of detail included in each style. In order to allow greater freedom in creating VR-visualizations that 1) correctly convey what is intended to be shown throughout the design process, and 2) have a varied and creative expression, it is important to be able to combine different styles. The differences and similarities in the approach towards visualizations between professional users with different backgrounds and/or different professional roles would be very interesting to study further. In this investigation we included architects and civil engineers, and found tendencies that we would like to investigate further with a larger number of participants. Since VR-visualizations are commonly used by city planners to communicate projects and proposals to the public, it would also be of great interest to include the general public as a new target group. Designing textures that will work in all the different stages of the design process is difficult, and this study must be considered as just a step in this process. To fully adjust and adapt the project to different design phases, a future aim is to incorporate it into a real design process, i.e. to follow a real planning project.

6. REFERENCES

References from Journals:
Al-Kodmany, K., 2002. Visualization tools and methods in community planning: From freehand sketches to virtual reality, *Journal of Planning Literature*, 17(2), pp. 189-211.

Brown, A. G. P., 2003. Visualization as a common design language: connecting art and science, *Automation in Construction*, 12, pp. 703-713.

Drettakis, G., Roussou, M., Reche, A. and Tsingos, N., 2007. Design and evaluation of a real-world virtual environment for architecture and urban planning, *Presence, Teleoperators and Virtual Environments,* 16(3), pp. 318-332.

Neto, P. L., 2003. Design communication: Traditional representation methods and computer visualization. *Visual Resources,* 19(3), pp. 195-213.

Savioja, L., Mantere, M., Olli, I., Äyräväinen, S., Gröhn, M. and Iso-Aho, J., 2003. Utilizing virtual environments in construction projects. *ITcon (Special Issue: Virtual Reality Technology in Architecture and Construction)*, 8, pp. 85-99.
Setareh, M., Bowman, D.A., Kalita, A., Gracey, M. and Lucas, J., 2005. Application of a virtual environment system in building sciences education. *Journal of architectural engineering*, 11(4), pp. 165-172

Unwin, S. 2007. Analysing architecture through drawing. *Building Research and Information*, 35(1), pp. 101-110.

References from Books:
Groat, L. and D. Wang, 2002. *Architectural Research Methods*. New York, John Wiley & Sons.

References from Other Literature:
Balakrishnan, B., Muramoto, K. and Kalisperis, L.N., 2007. Spatial Presence: An explication from an architectural point of view. *Proc. ACADIA Conference 2007: Expanding Bodies*, Halifax, Nova Scotia, Canada, Riverside Architectural Press, pp. 120-129.

Bassanino, M. N., 1999. The perception of computer generated architectural images. *PhD Thesis,* Liverpool, University of Liverpool.

Bermudez, J., 1995. Designing architectural experiences: Using computers to construct temporal 3D narratives. *Proc. ACADIA 1995, Computing in design, enabling, capturing and sharing ideas,* University of Washington, Seattle, USA, pp. 139-149.

Billger, M. and C. Dyrssen, 2005. Urban glades. Investigating light and sound in a research by design workshop. *Joining Forces. Design Research, Industries, and a New Interface for Competitiveness*, Helsinki, Finland.

Brown, A. G. P. and M. Nahab, 1996. Human interpretation of computer generated architectural images. *CAD and Creativity*, Bialystok, Poland.

Heldal, I., van Raalte, S. and Balot, K., 2005. Using virtual reality systems to support public participation in planning new roads. *Proc. the 11th International Conference on Virtual Systems and Multimedia*, *VSMM 2005*, Ghent, Belgium, pp. 76-89.

Karpenko, O. and J. F. Hughes, 2006, SmoothSketch: 3D free-form shapes from complex sketches, *Proceedings of ACM SIGGRAPH 2006*, Boston, Massachusetts, USA, pp. 589 - 598

Kwee, V., 2007. Architecture on digital Flatland - Opportunities for presenting architectural precedents. *Proc. ACADIA Conference 2007: Expanding Bodies*, Halifax, Nova Scotia, Canada, Riverside Architectural Press, pp. 110-119.

Lange, E., 2005. Issues and questions for research in communicating with public through visualizations. *Proc. Trends in Real-Time Landscape Visualization and Participation*, Wichmann Verlag, Heidelberg.

Oh J-Y., Stuerzlinger, W. and Danahy, J., 2006. SESAME: Towards better 3D conceptual design systems, *Proceedings of the 6th Conference on Designing Interactive systems,* University Park, PA, USA, pp. 80 - 89.

Panagiotis, P., Mania, K., Pletinckx, D. and White, M., 2006. Usability evaluation of the EPOCH multimodal user interface: designing 3D tangible interactions. *Proc. of VRST 2006*, Limassol, Cyprus, pp. 116-122.

Ritter, L., Li, W., Curless, B., Agrawala, M. and Salesin, D., 2006. Painting with texture. *17th Eurographics Symposium on Rendering 2006*, Nicosa, Cyprus.

Van Bakergem, W. D. and G. Obata, 1991. Free hand plotting - is it live or is it digital? *CAAD Futures '91*, Zurich, Friedrich Vieweg & Sohn Verlagsgesellschaft mbH, pp. 567-582.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

TOWARDS DIGITIZING COLOURS OF ARCHITECTURAL HERITAGE

W. Yan [a, *], P. Rajan [b]

[a] Dept. of Architecture, Texas A&M University, College Station, TX 77843, USA – wyan@archmail.tamu.edu
[b] Dept. of Computer Science, A&M University, College Station, TX 77843, USA – pankaj@cs.tamu.edu

KEY WORDS: Architectural Heritage, Colour, Digitizing, High Dynamic Range Imaging

ABSTRACT:

Architectural heritage represents the glorified past of the rich culture and architectural history of any country or community. It is important that the heritage be preserved and maintained for the generations to come so that they can learn from the heritage about their cultural identities as well as their architectural history. One of the essential approaches in preservation of architectural heritage is documentation of 3D geometries and surface textures of historical buildings. However, precise colour information, irrespective of lighting effects, which is an intrinsic property of the surface materials of building interiors and exteriors, has not been accurately documented in the scale of whole building surfaces, except for small sample areas. Every day the materials of historical buildings are decaying and their colours are fading. The goal of this project is to develop a method to assist in recording and documenting the intrinsic chromaticity information of building interiors and exteriors of architectural heritage with low cost and high efficiency. The method takes advantage of the emerging High Dynamic Range Imaging (HDRI) technology, which can store rich information about colour and illumination through digital photography. The significance of this research lies in that by recording the colour information, in addition to the geometry and texture information that can be obtained through other technologies, we can achieve more complete documentation for architectural heritage. In this paper, we present an overview of the problem and our effort of utilizing the advances in computer vision to document the colour information of architectural heritage.

1. INTRODUCTION

Architectural heritage represents the glorified past of the rich culture and architectural history of any country or community. It is important that the heritage be preserved and maintained for the generations to come so that they can learn from the heritage about their cultural identities as well as their architectural history. With the rapid industrialization, harmful substances are released into the atmosphere, which not only create health hazards but also have detrimental effects on architectural heritage. Besides pollution, the natural factors like earthquakes, hurricanes, and climatic changes are also affecting the beauty of historical architectural masterpieces. Though devising remedial measurements and restoration of the damage are the alternatives, it is equally important to document and monitor the condition of the heritage from time to time so that the damage can be accessed and worked upon at time to save considerable costs at the later time. While human inspection is the best way of diagnosis, it is time consuming and expensive. In this paper we will discuss some of our efforts towards utilizing the advances in computer vision technology to create an automatic system for documenting the colour information of architectural heritage.

One of the essential approaches in preservation of architectural heritage is documentation. With the advance of digital technologies, digital documentation, including digital photography, digital photogrammetry, and 3D laser scanning, etc., has been used successfully for documenting historical architecture. These digital technologies can assist in documenting the 3D geometries and surface textures of historical buildings. However, precise colour information, irrespective of external lighting conditions, which is an intrinsic property of the surface materials of building interiors and exteriors, has not been accurately documented, while colour information is one of the key identifying features of many architectural heritages. The colours have both symbolic and aesthetic significance, for example, in historical Chinese architecture, bold colour lacquers have been applied to structures. The colours of the roofs corresponded to the building purposes and ranks, e.g. yellow colours are restricted to the royalty only.

However, every day the materials of the historical buildings are decaying and their colours are fading. The colours we can see today will not be the same as the next generations can see if we cannot preserve the colour information. Figure 1 and 2 demonstrate the significant colour fading on the buildings ranging from the Forbidden City in Beijing, China to the Parthenon on the Acropolis at Athens, Greece.

The goal of the research presented in this paper is to develop an effective and efficient method to document the colour information, irrespective of external lighting conditions, of the interior and exterior surfaces of architectural heritage. The significance of the research lies in that by documenting the colour information, in addition to the geometry information, we can achieve more complete documentation about architectural heritage. The method can be used for detecting the colour change (e.g. colour decaying) of historical buildings if measurement is done in a regular basis, and for reconstruction of the historical building models with colour information of high fidelity.

Please note that throughout the paper, the term 'colour' we used refers to the intrinsic colour property of surfaces, excluding external lighting effect.

As digitizing colours of architectural heritage is a relatively new and large topic, the focus of this paper is an overview of the background, problem statement, related works, as well as our initial experiments towards the goal. We start in Section 2 by describing the background and problems of documenting colours for architectural heritage.

* Corresponding author.

2. BACKGROUND AND PROBLEM STATEMENT

Devices like spectrophotometers can be used for colour measurement in general, but they are limited to measure a small sample area (e.g. 3mm to 26mm). This makes it not applicable for measuring large areas like architectural interiors and exteriors. Thus, developing a technology to assist in documenting the colour information about the historical buildings becomes necessary.

Colour photography is able to record colour information of large surfaces to a certain extent. However, it cannot be applied directly to documenting precise intrinsic colour information of the surfaces. The reasons include: 1) in a colour photo, chromaticity (hue and saturation) and value of a colour on a surface cannot be separated from lighting effects and shadows; 2) the dynamic range of current photography is quite low, which basically means that a normal digital camera is unable to capture the colour details of the dark regions as well as very bright (or washed out) regions of a scene.

Figure 1. Three photos showing the colour decaying on historical buildings in the Forbidden City, Beijing, China. Top: a recent repainting (presumably with the original colours); middle: colours started to decay; and bottom: significant colour fading.

Figure 2. Two images showing the colour decaying on Parthenon on the Acropolis at Athens, Greece.
Top: Painting showing frieze slabs. (Artist: Alma-Tadema, Lawrence, 1836-1912. Birmingham Museum and Art Gallery).
Bottom: frieze of the current condition (Photographer: Athinaios, 2006, Creative Commons license)

In photography and imaging, dynamic range represents the ratio between the luminance of the brightest and the darkest regions in the scene. The range of luminance that human vision can perceive is very large, ranging from a sunlit scene of 100,000 cd/m^2 to starlight that is close to 0.001 cd/m^2, in which the ratio is hundred millions (Reinhard et al., 2006). A scene of the interior of a room with a sunlit view outside a window has a dynamic range of approximately 100,000:1 (Reinhard et al., 2006). However, the current normal digital cameras are capable of capturing the dynamic range of 255:1 and thus most of the colour details available in a scene are lost. Such imaging is called low dynamic range imaging (LDRI). As an example, let us consider the case shown in Figure 3. As it can be seen in the low dynamic range images of the interior of the church, depending on the exposures, either the windows can be washed out or the walls can appear very dark and their original colour information cannot be recovered from any single image (Figure 3).

Currently, there are different approaches of creating high dynamic range images. The most popular one, which is both high quality and low cost, is to combine photos of the same scene taken under different exposure levels (Reinhard et al. 2006). We use this method to create high dynamic range images in our project.

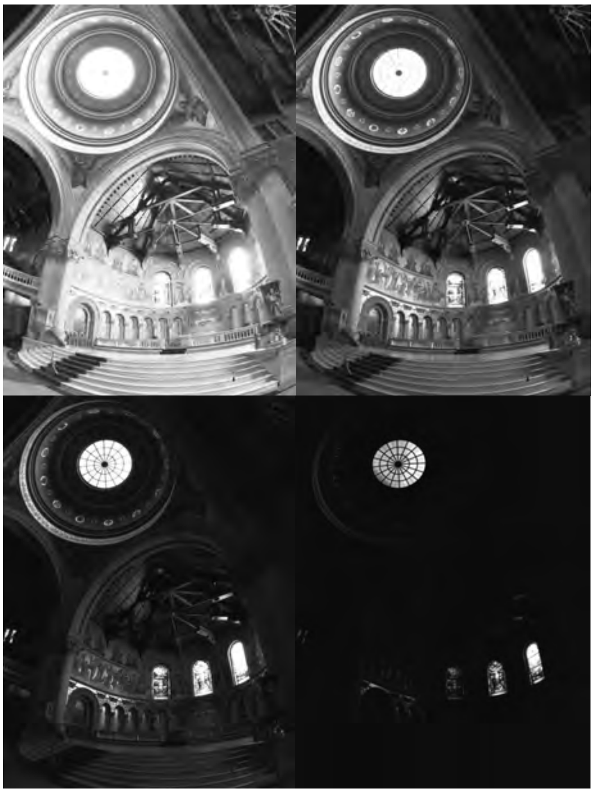


Figure 3. Photos of Stanford University Memorial Church taken at different exposure levels (Copyright Paul Debevec). Each photo is a low dynamic range image and the colour information is not well recorded by any single photo. For example, upper left: the ceiling structure is in good colour but the stained glass windows are washed out; upper right: the walls are good but the windows are still washed out and the ceiling structure is dark; lower left: the window on the left is good but both the ceiling and the walls are dark; and lower right: the middle window is good but the ceiling and the walls are too dark.

High dynamic range imaging (HDR/HDRI) has a clear advantage over LDRI because HDRI can record colour and illumination information in a much higher dynamic range. However, to be displayed on LDRI devices they need to be tone mapped, which means that the variations in the luminance represented by HDRI are remapped using a mapping function which utilizes the complete 255:1 dynamic range. Figure 4 shows a tone-mapped image of the Forbidden City.

Many algorithms are available for tone mapping (Reinhard et al., 2006). We have developed our own algorithm for tone mapping that is part of our colour documenting approach. Figure 5 shows our result of tone mapping for the Stanford Church photos.

From both Figure 4 and Figure 5, we can clearly see the rich colour information that can be recorded by HDRI, even though the images are tone mapped to LDRI for display purpose.

We will give a brief detail of some of the prominent work in HDRI tone mapping, as the starting point of our approach to recovering colour information that may be lost in the dark areas or washed out areas if using LDRI.

Figure 4. Tone mapped HDRI showing rich colour information. (source:http://www.flickr.com/photos/etherflyer/1393181113/in /photostream/)

Figure 5. Our tone-mapped HDRI using the photo set of the Stanford Church. The image preserved the colour information of the ceiling structure, the stained glass windows and the walls. This shows that HDRI will enable recovering the colour information of the scene.

Even with HDRI, shadows still present (Figure 4 and 5), which affects various vision-based computer algorithms dealing with techniques like image segmentation, recognition, as well as colour recovering. In our case, because of the presence of the shadows, the same surface colour will appear differently in well-lit areas and under shadows. As such, it is important that, we account for the presence of shadows in the digital images and try to recover colour information under the shadowed regions.

The effects of the shadows can be studied under two different categories that are cast shadows and form shadows (Figure 6). The *cast shadow* is the shadow formed when an object blocks the direct light. These are the common shadows like the shadow of a house on the ground when the sun is behind.

Another type of shadow is *form shadow*, which is caused by the shape of host objects and is also called self shadow or shading. Form shadows are shown on the object surfaces that are not directly facing the light and there are no cast shadows falling on the surfaces.

As we can observe in Figure 6, because of shadows, computer algorithms will be unable to tell in this image whether a region under shadow is of the same colour as its well-lit surrounding or of a different colour. For example the dark blue colour pointed by "Form Shadow" and its upper neighbour region's light blue colour pointed by "Well-lit Area" are actually the same colour in the real scene (they appear different because of lighting), but with only the information present in this single image, computers see them as two different colours (although humans can guess it correctly based on experience and knowledge about the buildings). In other words, computers will not be able to tell whether Figure 6 is a photo taken in an actual scene (where the appeared two blue colours are of the same lacquer), or it is a photo of another photo of the scene (where the two blue colours are of two different inks). Figure 7 further shows another sample image that can fool a human: a photo of Eric Grohe's mural, in which different shades of the green bridge are actually painted using different colour acrylics. Our point here is that our computer algorithms should be able to tell that Figure 6 is an actual photo of a scene (therefore conclude the two blue shades are of the same lacquer) and Figure 7 is a photo of a mural (therefore conclude the green shades are of different acrylics). Of course, we need more information than just a single image for each case.

Our approach to the shadow problem is utilizing HDRI and time-varying (therefore lighting condition varying) images, as will be described in Section 4, prior to which, related works will be discussed in Section 3.

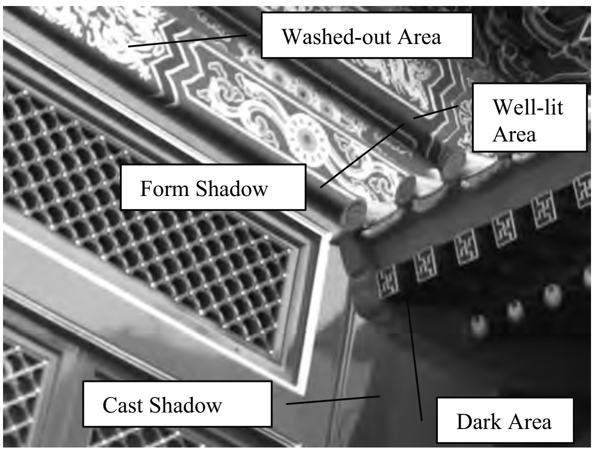


Figure 6. Image showing cast shadow, form shadow, well-lit area, dark area, and washed out area.

Figure 7. Eric Grohe's mural: acrylic on cement block, 34' x 130', Bucyrus, Ohio (Source:http://www.ericgrohemurals .com)

3. RELATED WORK

Research has been conducted to recover the reflectance property of surfaces using images of low dynamic range (e.g. Weiss, 2001 and Tappen et al. 2005). Similar methods have been investigated for preserving cultural assets (Nishino et al. 2002). However, there are limitations of using LDRI as by nature it doesn't contain enough information about colour, and the performance of these algorithms depends heavily on the lighting conditions. In historical building sites, lighting can hardly be controlled and this makes it hard for those algorithms to perform as expected. Another recent approach to recover the colour information of architecture exteriors was developed by Tchou et al. (2004) and Debevec (2005). This approach can recover surface colours, but it is complicated because it uses a model of the scene's geometry, a set of photographs of the scene taken under natural illumination, and corresponding measurements of the illumination. It also employs an iterative inverse global illumination technique to compute the surface colours. Therefore it cannot be efficiently applied to preserving the colour information of interior and exterior surfaces of architectural heritage.

Our approach to recovering colour of architectural heritage is based on the related work of HDRI, its tone mapping, and shadow removal.

3.1 Related Work in HDRI and Tone mapping

HDRI requires that images be tone mapped so that they can be displayed in LDRI devices. Tone mapping presents in the workflow of our method for shadow removal and the display of image processing results, though the mapping doesn't provide the actual colour values we are aimed to obtain.

Various tone mapping operators have been researched and discussed in the literature. The tone mapping algorithms have been divided into two categories, local and global operators. Some of the prominent global operators used are that of Tumblin and Rushmeier (1993), Ward et al. (1997), Drago (2003) and Rienhard (2004). Tone mapping operator by Tumblin and Rushmeier et al. (1993) employs the concept of brightness preservation and perception using the psychophysical model of brightness perception proposed by Stevens and Stevens (1960). Ward (1997) proposed a method based on the histogram equalization in which the dynamic range is allocated to the pixels proportional to the amount of pixels having that brightness. Drago (2003) based his logarithmic operator for tone mapping on the observation that the human visual system uses the logarithmic response to intensities. Reinhard et al. (2004) based their operator on the logarithmic response of the eye to intensities and developed a photoreceptor model for the compression.

Among the local operators, Rahman et al. (1996) exploited the Retinex theory proposed by Land (1971) to develop a tone mapping operator based on the local luminance properties of the scene rather than the global luminance. Pattnaik et al. (1998) introduced the multi-scale observer model which introduced the concept of centre surround processing which takes into account the representation of pattern, luminance and colour processing in human visual system. Reinhard et al. proposed a local operator based on the zone system (2002). Another prominent algorithm based on the local characteristics of the image was given by Fattal (2002). In Fattal's algorithm, the magnitude of high gradients in the gradient field of the luminance image is attenuated for obtaining the low dynamic range image. This method is able to preserve the fine details. Besides, this method is also able to avoid some of the major problems in the tone mapping technique like halo formation and loss of local contrast.

Almost all of above algorithms derive their inspiration from the human eye perception, while our approach is simple and aimed at assisting the retrieval of physical values of surface colours.

3.2 Related work in Shadow Removal

We have investigated some of the prominent research work in the field of shadow removal. Most of the shadow removal algorithms derive their roots from the concept of intrinsic images introduced by Barrow and Tenenbaum (1978) according to which, an image I can be defined as the product of a reflectance R and luminance image L, mathematically:

$$I(x, y) = R(x, y) \times L(x, y) \qquad (1)$$

This work was exploited by Weiss (2001), who was able to extract the luminance and the reflectance components by reconstructing the median of the log derivative of the image sequence taken over time.

Finlayson and Hordley (2001) derived a 1-dimensional illuminant-invariant shadow free image, by projecting the 1-D chromaticity information in the direction orthogonal to the direction of lighting. The reflectance image is reconstructed by comparing edges in the 1-D image intrinsic image to the edges present in the actual intensity image. In Baba et al. (2004), an algorithm based on shadow density has been proposed, which can eliminate the shadow from a single texture image. In Tappen et al. (2005), intrinsic reflectance image is derived from a single image using a classifier, which is trained to classify illumination and reflectance change by using a training set consisting of the derivative information of the illumination change and the reflectance change. Levine and Bhattacharyya (2005) presents a machine learning approach based on support vector machines which are trained to classify the colour ratios of the shadowed and the non-shadowed regions.

Most of the above algorithms are not applicable to our problem because of the restrictions they comply. Our shadow removal, however, was mostly inspired by Weiss (2001) work mentioned above that deals with cast shadow situation using time-varying image sequences, whereas we use both time-varying and exposure-varying images for both cast shadow and form shadow removal. The idea is to collect as much information as possible to produce a simple solution that can be used efficiently to document colour information of architectural heritage.

4. SOLUTIONS AND METHODOLOGY

Our first phase is to obtain colour images excluding lighting effects – there will be no cast or form shadows, and no colour loss caused by dark or washed out areas. Each pixel of a resulting image represents a small sample area on a surface. Further, based on the image, a physical colour value, such as CIE L*u*v*, with the calibration of a spectrophotometer, will be added to each pixel or sample.

We are able to obtain the HDRI radiance map (brightness values of a complete dynamic range) by combining images of different exposures using the method proposed by Debevec and Malik (1997). HDRI can eliminate the problem of the loss of colour information in the dark areas and the washed out areas. Combining time-varying features, we are able to develop algorithms to deal with the cast and form shadow problems.

4.1 Cast shadow removal

For dealing with cast shadow, our approach utilizes the advantage of the information about the dynamic range, conveyed by multiple exposures, as well as the shadow movements to identify the shadow edges. This method is based on Weiss (2001) by using the shadow movement information obtained from time-varying images, but we added a new dimension of useful information – exposure-varying images (HDRI) – to attack the shadow removal problem and make the algorithm potentially applicable to real world problems such as the one we are targeting: digitizing colour information of architectural heritage. Figure 8 shows the test images taken at different times and with different exposures for a simple scene with a Lego. We have developed two algorithms for cast shadow removal that are similar in the following steps: pre-processing to align all the images, identifying shadow and objects in the images, brightening the shadow region. The difference of the two algorithms is: while one algorithm uses a similar reconstruction method like that proposed in Weiss (2001) with time-varying images, the other replaces the shadow pixels with the non-shadow pixels from images of different time settings (therefore different lightings). The details of the algorithms are being described in a separate paper. Figure 9 shows the result of our first algorithm for cast shadow removal.

4.2 Form shadow removal

We have developed an algorithm that deals with form shadow removal by examining the hue constancy, region connectivity, and radiance change over time among segmented colour regions including both shadow and well-lit regions. We also use both exposure and time varying images in order to utilize the high dynamic range of colour information and the lighting change over time on the scene objects. The latter makes the radiance to change differently for surfaces facing different directions, and this difference can be used for detecting whether an intensity difference between two regions with the same hue is caused by the lighting difference or by the material difference. This allows us to identify the intrinsic colours of the surfaces.

Our algorithm can be explained using the pseudo code below:

Given time and exposure varying images of the same scene (Figure 10):

1. *Obtain an HDR tone-mapped image for the exposure-varying image sequence at each time setting. Say we get A, B, C, and D four images.*
2. *For each of A, B, C, or D, do segmentation and labelling based on Hue/Saturation/Value (using the method such as Felzenszwalb et al. 2004). Each image may be segmented to a different number of regions. Merge the four segmented images to get F that has all segmented regions labelled (the number of segmented regions in F may be larger than that of each single image of A, B, C, or D).*

   This may introduce more labels than the actual colours, e.g. two regions of the same colour may have different values due to lighting difference (and labelled differently). Therefore we need to check further:

3. *Connectivity. Find all connected segment pairs on F.*

4. *For each connected segment pair Si and Sj on F*
   If Si and Sj have the same Hue
   If Si and Sj Radiances change over time differently
   Si and Sj have high probability of being the same colour. Given architectural or many other cases, this basically means they are the same. They will be labelled the same.
   Else
   Si and Sj must be different colours.
   Else
   Si and Sj must be different colours.

5. *Using the re-labelled image F, recolor the image. That way we can display an image that is segmented based on colours irrespective to lighting (Figure 11).*

Basically, assuming architectural images, we expect that if two colour regions are connected (say one is in a well-lit region and the other is in adjacent form shadow region), having the same hue, but their changes of radiance values over time are different (caused by lighting changes), the two regions will be labelled as having the same colour, and therefore the form shadow effect is removed. Figure 10 and 11 show the input and output images of the algorithm. This algorithm will be able to deal with the case of the mural in Figure 7 – given a set of time- and exposure-varying photos about the mural, any two connected but segmented regions will have the same radiance change over time, as they are facing the same direction, therefore they will be labelled as different colours, which is true because they are painted with different acrylics.

5. CONCLUSIONS AND FUTURE WORK

The paper presents our research towards digitizing colours of architectural heritage, including an overview of the subject and our initial results of image processing methods that can be used further in case studies. Our future work includes: integrating cast shadow and form shadow removal; assigning physical colour measurements, such as CIE L*u*v*, with the calibration of a spectrophotometer, to pixels representing surface samples; designing a database for efficiently storing calculated colour information for buildings; case studies in selected historical buildings; and investigating the integration of digitizing colours and digitizing geometries for architectural heritage.

Acknowledgements

This research was supported by the National Endowment for Humanities (NEH) Digital Humanities Start-Up Grant.

6. REFERENCES

Baba, M., Mukunoki, M., and Asada, N., 2004. Shadow removal from a real image based on shadow density. In *ACM SIGGRAPH* Posters.

Barrow, H.G. and Tenenbaum, J.M., 1978. Recovering intrinsic scene characteristics from images". In A.Hanson and E. Riesman, editors, *Computer Vision Systems*. Academic Press.

Debevec, P., Malik, J., 1997. Recovering High Dynamic Range Radiance Maps from Photographs, *Computer Graphics (Proceedings of SIGGRAPH 97)*, ACM.

Debevec, P., 2005. Photometric Realism: Accurate Reflectance Modeling of Sites and Artifacts, *VAST*.

Drago, F., Myszkowski, K., Annen, T. and Chiba, N., 2003. Adaptive Logarithmic Mapping For Displaying HighContrast Scenes. *EUROGRAOHICS* Vol 3.

Fattal, R., Lischinski, D. and Werman, M., 2002. Gradient Domain High Dynamic Range Compression, in *ACM Transactions on Graphics*, 21, pp. 249–256, July 2002.

Felzenszwalb, P. F. and Huttenlocher, D. P., 2004. Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision*, Vol. 59, No. 2, September 2004.

Finlayson, G.D and Hordley, S.F., 2001. Color constancy at a Pixel, *J.Opt Soc Am. A*, 18 (2):253-264, Feb 2001.

Land, E. H. and McCann, J. J., "Lightness and the Retinex Theory," Journal of the Optical Society of America 61(1).

Levine, M.D. and Bhattacharyya, J., 2005. Removing shadows, *Pattern Recognition Letters* 26, 251-265.

Nishino, K., Hara, K., Tan, R., Miyazaki, D., Ikeuchi, K., 2002. Photometric Aspects on the Preservation of Cultural Assets, in *Proc. of Virtual Systems and Multimedia (VSMM'02)*, pp.926-933, September 2002, Korea.

Rahman, Z., Jobson, D. J. and Woodell, G. A., 1996. A Multiscale Retinex for Color Rendition and Dynamic Range Compression, in *SPIE Proceedings: Applications of Digital Image Processing XIX*, Vol. 2847, Denver, CO: SPIE.

Reinhard, E., Stark, M., Shirley, P., and Ferwerda, J., 2002. Photographic tone reproduction for digital images. *Proceedings of SIGGRAPH* 2002.

Reinhard, E. and Devlin, K., 2004. Dynamic Range Reduction Inspired by Photoreceptor Physiology, in *IEEE Transactions on Visualization and Computer Graphics*.

Reinhard, E., Ward, G., Pattanaik, S., Debevec, P., 2006. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, Elsevier.

Stevens, S. and Stevens, J., 1960. Brightness Function: Parametric Effects of Adaptation and Contrast, *Journal of the Optical Society of America* 50, p. 1139A, Nov.

Tappen, M., Freeman, W., and Adelson, E. 2005. Recovering Intrinsic Images from a Single Image. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 9, pp. 1459-1472, September, 2005.

Tchou, C, Stumpfel, J., Einarsson, P., Fajardo, M., and Debevec, P. 2004. Unlighting the Parthenon, *SIGGRAPH Sketch*.

Tumblin, J. and Rushmeier, H., 1993. Tone Reproduction for Computer Generated Images, *IEEE Computer Graphics and Applications*, 13(6):42–48, November 1993.

Ward, G., Rushmeier, H., and Piatko, C. 1997. A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes, *IEEE Transactions on Visualization and Computer Graphics*, 3(4).

Weiss, Y., 2001. Deriving intrinsic images from image sequences, in *Proc. of 9th IEEE Int'l. Conf. on Computer Vision*, 68–75, Jul., 2001.

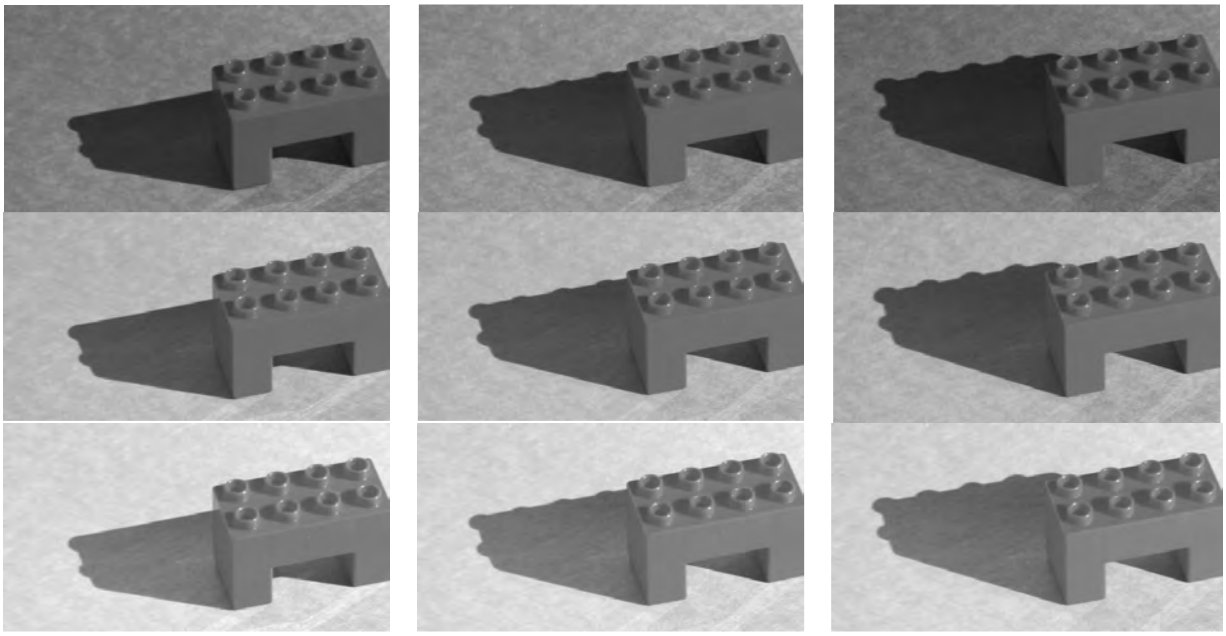


Figure 8. Time and exposure varying images: three sets (columns) of images taken at different times 20 minutes apart from each other, in which shadow movement can be seen. Multiple exposures (rows) were used for each set of the images (only 3 out of 13 different exposures are shown for each set of images).

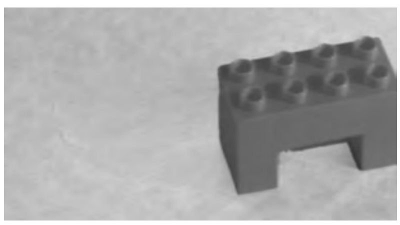


Figure 9. Our result obtained for cast shadow removal.

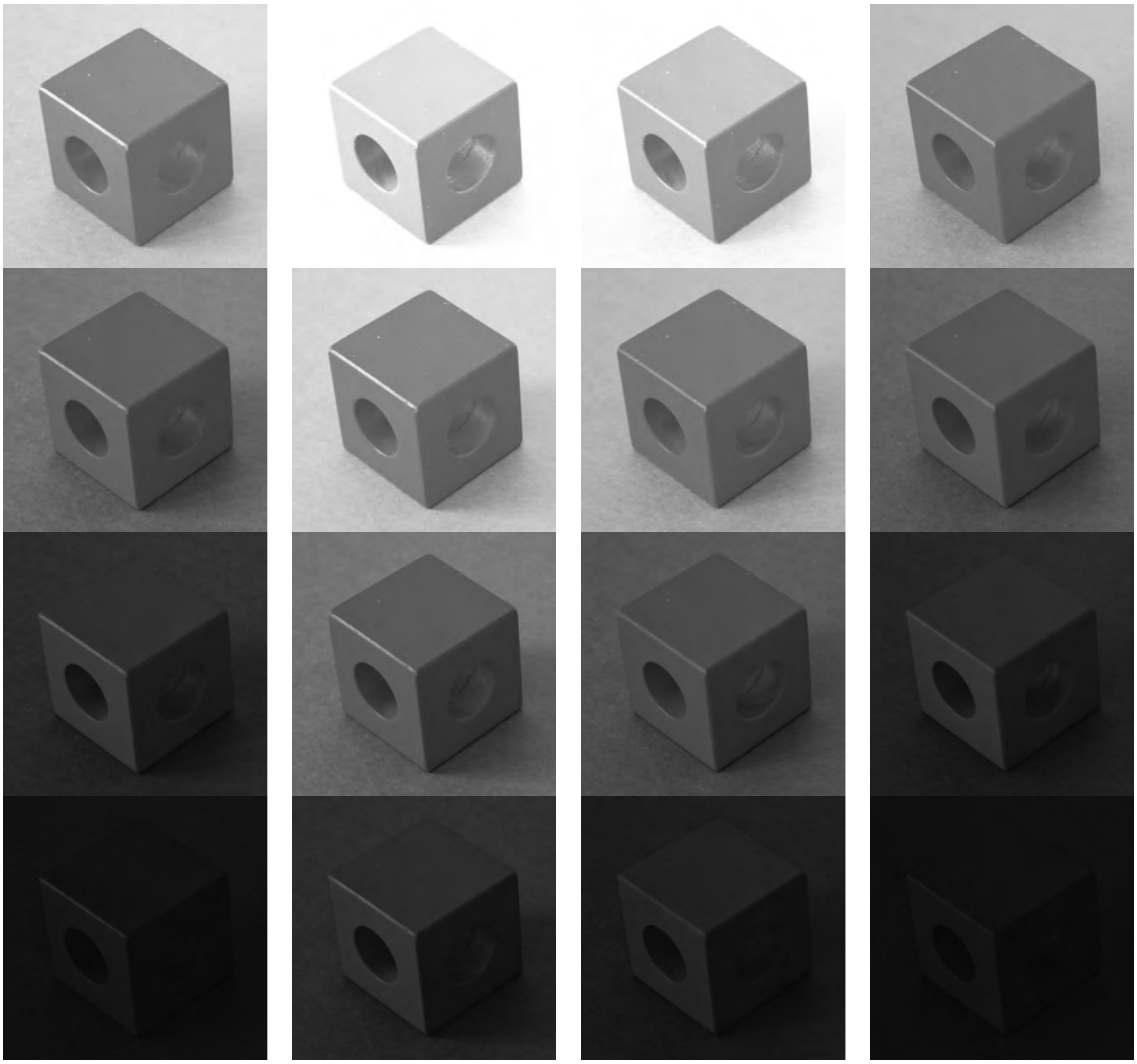


Figure 10. Time- and exposure-varying images: four sets (columns) of images taken at different times, 1-3 hours apart from each other during a day, in which lighting difference can be seen. The radiances of surfaces facing differently changed differently over time. Multiple exposures (rows) were used for each set of the images (only 4 out of 22 different exposures are shown for each set).

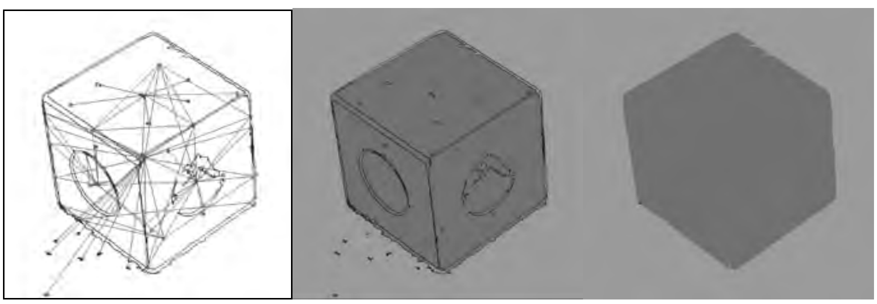


Figure 11. Left: segmented regions and their connectivity shown by the lines connecting region centroids. Note that the centroid of the background is at about the centre of the entire image and has maximum number of connections with other centroids. Some very small regions are formed due to their difference in brightness against their surrounding areas. Middle: each region, whose boundaries are shown, is labelled correctly by our algorithm. Right: result without region boundaries. The choosing of the colours present in the above middle and right results is arbitrary as later actual physical colour values will be assigned to each region based on the measurement of spectrophotometers.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

FROM INTERNET-BASED MULTIMEDIA DATA COLLECTION TO 3D VISUALIZATION OF VIRTUAL UNDERWATER SITES

N. Cheaib [a, *], S. Otmane [a], M. Mallem [a], A. Nisan [b], J.M. Boi [b]
[a] IBISC CNRS FRE 2873, University of Evry 91020 Evry, France
(nader.cheaib, samir.otmane, malik.mallem)@ibisc.univ-evry.fr
[b] LSIS UMR CNRS 6168, ESIL 163 Luminy Avenue, 13288 Marseille France
(adrien.nisan, jean-marc.boi)@univmed.fr

KEY WORDS: Groupware, Web Services, Software Agents, Tailorability, 3D visualization and modelization

ABSTRACT:

In this paper, we present a two-fold research work in the context of a national project, *DIGITAL OCEAN* for the creation and distribution of multimedia content. In the first part, we present tailorable groupware architecture, Oce@nyd, to allow a collaborative collection of multimedia information over the internet. The Oce@nyd software architecture is based on Web services and software agents' integration. This research is, until now, never been exploited in the context of groupware tailorability, hence bringing innovation in the CSCW (Computer Supported Cooperative Work) domain. The second research part concerns a 3D underwater viewer, *Oceanyd*. In this scope, some questions have been raised concerning terrain rendering with excavations like overhangs, holes or caves. In fact, the Digital Elevation Model (DEM), provided by Semantic-TS, a partner of the project, does not allow this kind of modelization. Hence, approximating the terrain by a finite set of authorized cases is one of the suggestions raised. Following this restriction, generating the different Levels-Of-Detail (LOD) which will be used during the real-time 3D visualization would be a solution and this by generalizing the *GeoClipmaps* concept from *Losasso and Hoppe.*

1. INTRODUCTION

With the advancement in computer entertainment and the apparition of Internet technologies, universal interoperability within collaborative applications is becoming a reality, while users geographically distributed people are highlighting the flexibility of cooperation by exchanging universally accessible services but often using incompatible applications that may lead to interoperability problems (Cheaib, 2008). As a result, the need for more effective means of collective collaboration is an extremely valuable area of research, while the demand increases for a framework to enable users to interact and collaborate based around mutual goals and shared data. However, most of these systems do not take in consideration the evolving need of users' to dynamically integrate new functionalities in order to enhance collaboration with others. In fact, making the system, its interfaces and the services that they could offer tailorable for users, is an essential and an ongoing research field that needs much attention to yet be concrete (Cheaib, 2008). For this reason, tailorability has shown to be an essential property that should be taken in consideration, as it offers to users the possibility to adapt the application based on their needs and preferences, and not the other way around.

On one hand, Web services have become one of the most important architectures for the cooperation of heterogeneous systems, and have ushered in a new era of software design that focuses on implicit and explicit collaboration between organizations (Gannod, 2007). While computer networks have been able to pass data between different hosts, it was the emergence of Web services that allowed these remote hosts to offer services in a more flexible and dynamic way. On the other hand, an important benefit in the use of software agents in designing software and groupware applications is their ability to help through collaboration, human beings and software logistics, while their concept is even older than Web services and has been used successfully for the implementation of distributed applications. In this paper, we present an innovative approach for tailorable groupware architecture, Oce@nyd, which is based on Web services and software agents. Our groupware model is conceived in order to enable users to collaboratively share multimedia files on the internet, while offering means to tailor the services in the system by the use of software agents that will discover external Web services and integrate them into the application, hence offering new functionalities without affecting users' collaboration, neither system's execution.

1.1 Tailorability and need of new Architecture

Some definitions exist in the literature for the concept of tailorability, but it is still ambiguous in putting it forward in the CSCW domain, where technologies for implementing such a concept are still not explicitly identified. We retained few definitions that seemed most interesting to our work, as in (Stiemerling, 1998) that defines a tailorable application as a system that can be adapted properly according to changes and the diversity of users' needs, or (Slagter, 2001) that defines tailorability as the capacity of an information system to allow a person to adjust the application based on personal preferences or different tasks. For (Foukarakis**,** 2007), tailoring is the continued development of an application by making persistent modifications to it. It is in fact initiated in response to an application being inefficient to use. In our work, we define tailorability in groupware as follows: "a tailorable groupware is a collaborative system that can be dynamically (dynamic integration of new functionalities with minimal human assistance) or statically (the user can explicitly add functionalities to the groupware by extending its code) adapted to satisfy users' preferences". However it remains to determine the mechanisms of evolution of tailorability. Morch (Morch, 1997) defines the concept in terms of customization, integration and extension. Tailorability by customization is limited by a set of predetermined number of components, tailorability by integration is to insert a new component in the architecture of the application, and tailorability by extension or radical tailoring is offering means to change or extend the components' implementation in order to derive the same flexibility as an "initial" application design. These mechanisms offer more flexibility but require more and more from the user computer skills, which partly explains why most of the current CSCW systems generally steer their tailorability to developers and expert users rather than end users that, paradoxically and from social sciences and humanities research, are those who need it the most. In this article, we focus on the third type of tailorability, hence extending program code by new components that are, in our work, Web services in order to satisfy users' preferences. Hence, we assume that a component is a Web service initiated by an agent constituting the system, which takes the initiative on behalf of the user to search for Web services having specific functionalities and provide them as agents' services. In groupware, a mismatch between the task done by users and the corresponding technology they are using could affect the co-operating people (Slagter, 2003), thus tailoring by end-users themselves is generally regarded as a suitable mean to solve this problem. Due to a lack of a theoretical framework for tailorability and the corresponding evaluation methods, results of different studies for groupware tailorability are hard to compare. In fact, our research is mainly concentrating upon:

• Development of a collaborative architecture supporting tailorability.

• Integration of technologies that has not been exploited before in the context of groupware tailorability.

We will proceed as follows; in the second section we give a background on the concept of collaboration, along with Web services and software agents, while we explain the reason behind choosing these technologies as basic components in our conceptual model, which constitutes the research part of the project we are working on. In section 3, we explain our groupware model, Oce@nyd, which is based on Web services and software agents. In section 4, we present a case study of the conceptual model applied on the DIGITAL OCEAN project. In section 5, we will briefly talk about research concerns of the Oceanyd application, a 3D sea-bed view that is a part of the global project. Finally, we give a conclusion and perspectives in the field.

2. BACKGROUND

2.1 Ellis's 3C model

We refer to the 3C model (Ellis, 1991) shown in Figure 1 for further understanding of the term collaboration and the functionalities behind it. In fact, a groupware system covers three domain specific functions, production/cooperation, communication and coordination, as shown in Figure 1: The production space designates the objects resulting from the activity of the group (ex: word document, paint etc.). For Ellis (Ellis, 1991), this production space is concerned with the result of the common tasks to be achieved and it is the space where the productivity will take place. The coordination space defines the actors and their social structure, as well as the different tasks to be accomplished in order to produce objects in the production space. Ellis eventually completed the model with the communication space that offers for the actors in the coordination space means to exchange information in which the semantics concern exclusively the actor, and where the system only acts as a messenger.

Figure 1: 3C Model

In this article, we will use this decomposition of groupware's functionalities in order to introduce a collaborative architecture supporting the functional decomposition of services that can be present in a groupware system.

2.2 Web Services and the World Wide Web

W3C (W3C, 1994)) defines a Web service as follows: "It is a software system that acts as an interoperable support in the machine-machine interaction. In fact, the service-oriented architecture (SOA) emerged due to its simplicity, clarity and normalized foundations. The concept of Web Services currently revolves around three acronyms (Newcomer, 2002) as we can see in Figure 2:

• **SOAP** (Simple Object Access Protocol) is a protocol for inter-application exchanging that is independent of any platform and based on XML. A SOAP service call is an ASCII flow embedded in XML tags and transported to the HTTP protocol.

• **WSDL** (Web Services Description Language) gives the XML description of Web services by specifying the methods that can be invoked, their signatures and access point (URL, port, etc.). It is therefore equivalent in a way to the IDL language for CORBA distributed programming.

• **UDDI** (Universal Description, Discovery and Integration) is a standard of a distributed directory of Web Services, allowing both publishing and exploration. UDDI acts as a Web service itself, whose methods are called using the SOAP protocol.

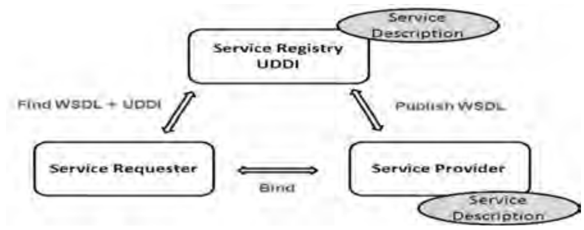


Figure 2: Web Services Scenario

Our choice of using Web services in our system is driven by the fact they are: Language and platform independent (separation of specification from implementation), deployed over the Internet (no centralized control, use of established protocols), loosely coupled (using synchronous and asynchronous interactions.) and interoperable (using standards already deployed and functional to support systems interoperability).

2.3 Software Agents

There exist several definitions of software agents in the literature. The authors in (Khezami, 2005) have identified the agent as a computing object (in the sense of object-oriented languages) whose behavior can be described by a script with its own means of calculation, and can move from a place to another in order to communicate with other agents. According to (Maamar, 2003), an agent is a piece of software that acts on an autonomous basis to initiate missions on behalf of users. In fact, agent technology is considered nowadays one of the most interesting technologies to support SOA, and is well suited for applications that are communication-centric, while software agents have the potential to dynamically discover services or behaviors that are, in our paper, based on Web services, and which can be processed concurrently.

2.3.1 JADE Platform: Java Agent DEvelopment framework (JADE) (Tilab, 2002) is a middleware written in Java that simplifies the development of software agents by providing basic services as well as a set of tools for deployment. The platform contains a runtime environment where JADE agents may evolve while being active on a given host, a library of classes used to develop agents and a suite of graphical tools that allow the administration and supervision of agents' activities at runtime.
The main container contains two special agents:
• **AMS** (Agent Management System) which provides a service Namespace (i.e. it ensures that every agent in the platform has a unique name) and represents the authority in the platform (it is possible to create or kill agents in remote containers by calling the AMS).
• **DF** (Directory Facilitator), on the other hand, is analogous to the UDDI used by Web services, and offers the Yellow Pages service through which an agent can find other agents providing the services it needs in order to achieve its goal. JADE defines a generic agent model that can perform any type of architecture while fully integrating the FIPA (FIPA, 2008) communication model: interaction protocols, wrapping, ACL (Agent Communication Language), languages content and transport protocols. In what follows, we proceed with some related work exposing researchers' motivation in this domain that affects many applications areas.

3. OCE@NYD GROUPWARE MODEL

As mentioned before, the aim behind our model is twofold: First, we want to integrate software agents and Web services, which were unfortunately, implemented using different standards, into a cohesive entity that attempts to surpass the weakness of each technology, while reinforcing their individual advantages (Shen, 2007) in the context of tailorable groupware design. The second aim is to bring our research results into practice and this by using our conceptual model on a real life project, thus shifting groupware development from design to real practice.
We extend the work in (Cheaib, 2008) for the use of SOA in the design of tailorable groupware, as it offers the needed interoperability and reconfigurability between system components, and the importance of using software agents in order to enhance the discovery of web services by making them proactive and dynamic. Moreover, we rely on the Arch model (Bass, 1992) by offering a canonical decomposition of the main structure of the system into five main components (Functional core, Functional core adapter, Physical Interaction, Logical Interaction and Dialog Controller components), each having a specific functionality. However, in this article we will concentrate on the design of the functional core (FC) which is the main component of the system along, and we make no assumption about the other components.
We can see the global architecture in Figure 3 below. In fact, we rely also on Dewan's conceptual model (Dewan, 1998) that structures a groupware system into a variable number of layers, each representing specific levels of abstraction, where the highest layer is the semantic layer that corresponds to the functional core of the system (coincides with the one of the Arch model), and the lowest layer representing the material or the hardware level (Arch's Physical Interaction component), and eventually we compare our model with the one in (Laurillau, 2002) that is itself built using the later models. However, the Functional Core (FC) represents the two shared layers of the architecture as we can see in Figure 3, and enables all users to manipulate domain objects and have access to various services during the interaction with the system, which is different from the clover model (Laurillau, 2002) that splits the FC into one shared layer and one replicated for every user. In our model, all other layers besides the FC are replicated, and handle the set of services and the state of the system that is private for every user in collaboration.

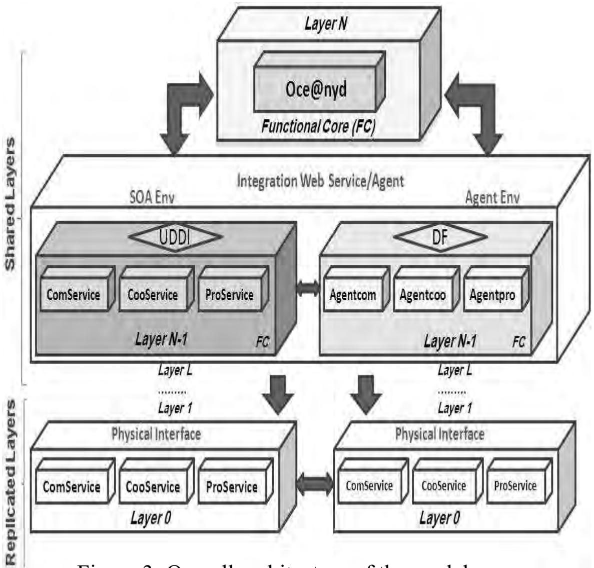


Figure 3: Overall architecture of the model

Hence, the aim of the FC in our model is for:
• Software agents to discover invoke and publish Web services in the UDDI.
• Web Service clients to discover and invoke software agents' services in the Directory Facilitator (DF) of the JADE environment.
• Web Services to be published in the Directory Facilitator (DF) as agent services.

Note that Figure 3 representing our proposed architecture shows only the FC of the system along with the physical interaction layer that implements the interactions with the users. In the last section, we implement the physical component as a Web interface serving as a case study of our model.

3.1 FC Decomposition

As we can see in Figure 3, the first component on the level N-1 is based on a SOA environment, and contains all the Web services in the system grouped into three main services, satisfying Ellis' 3C model (Ellis, 1991): communication services, coordination services and production services.
• **ComService:** contains all services offering means of communication between users in collaboration (videoconference service, voice recorder service etc.).
• **CoorService:** contains services implementing rules of coordination by codifying their interaction (i.e. workflow services).

• **ProService:** contains services that are the collaborative product of using the architecture. (Ex: Paint application, Word document etc.).

These services' classes can be considered as orchestrations of various services in the system (Peltz, 2003), and include services based on the functionalities they offer.

In parallel to the SOA environment, a JADE environment constitutes the other part of the FC on the level N-1, where, as in the SOA environment, agents are grouped into three main classes: communication, coordination and production agents. The functional decomposition of the layer into three main sub-components will enhance the interaction with Web services in the system, while every agent in one particular sub-component would be bound by a Web service of the same class (communication, coordination or production), in order to fasten the integration and deployment's process by insuring system's modularity. Therefore, each sub-component in this layer manipulates semantic objects dedicated to one of the 3C model functionalities, and performs specific processing functions on its services. We suppose, for now, that only the layer on the level N-1 of the FC satisfies the three main classifications according to the 3C model, while we have made no assumption about the decomposition of the highest semantic layer of the FC on the level N, which is composed of one single component with a set of codec for integrating Web services with agents, as we will see in the next section.

3.2 FC Implementation

As we have mentioned, the FC on the level N shown in Figure 3, should allow mechanisms for translating Web services' invocations into a language understood by software agents, and vice versa. Few related work in the literature have been identified dealing with such translation mechanisms. The authors in (Nguyen, 2005) present their tool, WS2JADE, that is based on two distinct layers: An interconnection layer that glues agents and Web services together, and a static management layer that creates and controls these interconnection entities called WSAG (Web Service Agents), that are able to communicate and deliver Web services as their own services by producing and deploying WSAG at runtime. For (Maximilien, 2004) agents can represent service consumers and providers that are independent, and collaborate together in order to dynamically configure and reconfigure service-based applications. Their approach implements an agent-based architecture and is realized in a Web Service agent platform (WSAF) that uses QoS ontology and an XML language enabling consumers and service providers to expose their preferences.

Our primary concern is to have a tool translating Web services' and agents' invocation messages, and creating interaction mechanisms in order to tailor Web services in the system. In fact, a useful tool in using JADE is the WSIG ('Web Service Integration Gateway') (Board, 2005) that meets our needs by providing means to register Web services in the JADE DF "mapped" with descriptions of FIPA agents. In this case, registered Web services can be called by agents by directing the invocation to the WSIG. Thus, a Web service is published as JADE agent service, and an agent service can be symmetrically published as an "end-point" of a Web service.

As shown in Figure 4, the highest level of the FC contains a WSIG consisting of several components, each linked either directly or indirectly to two registries, the DF that is not visible outside the platform, and the UDDI is visible internally to the WSIG, and externally to Web services and Web service clients, but not directly to agents, and hence the WSIG acts as the interface point between the agents and the SOA environment.

In order to be visible in both environments, WSIG is registered as a special agent service in FIPA DF (Directory Facilitator) and a special Web service endpoint in the UDDI directory. The purpose is to ensure that any registered Web service is visible to agents via the DF and any registered agent is visible to Web service clients via the UDDI. The WSIG is implemented as a standard JADE agent with behaviours controlling access to a number of local components:

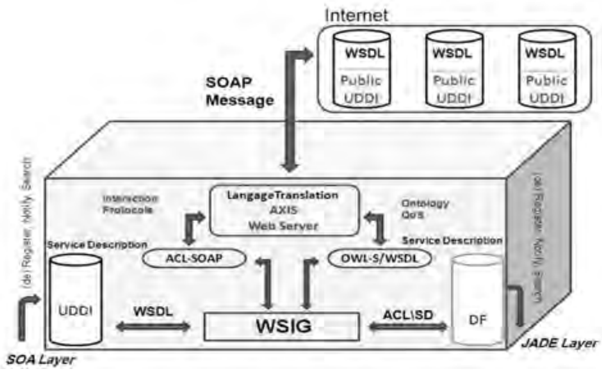


Figure 4: Highest semantic layer in the model

• **OWL-S/WSDL** codec that has the purpose of generating OWL (W3C, 2004) ontologies from WSDL specifications. The aim is to convert Web services' data types and operation inputs and outputs into agent ontologies, and thus into meaningful information for agents as they need to know how to invoke operations of a Web service.

• **ACL/SOAP** codec is responsible for parsing ACL messages received from the DF in order to extract the encoded service descriptions (SD) held within their content, then translating the SD into a web service invocation and returning the results to the WSIG for registration in the UDDI. This codec also operates in a bidirectional manner in order to translate SOAP and service specification information into correctly encoded ACL messages and DF entries.

• **Axis' JAX-RPC** (Java API for XML based Remote Procedure Call) is an application program interface (API) that enables Java developers to include remote procedure calls (RPCs) with any Web-based applications. It is aimed at making it easier for applications or Web services to call other applications. The JAX-RPC programming model simplifies the development by abstracting SOAP protocol-level runtime mechanisms and providing mapping services between Java and the Web Services Description Language (WSDL).

As we can see in Figure 4, language translation between the two environments for making the discovery and integration of new services in the system dynamic is leveraged by reusing and identifying existing technologies instead of reinventing the wheel. In the next section we will see how our model can be put into practice.

4. CASE STUDY- OCE@NYD

Figure 5 illustrates the application of the model on an ongoing project in our laboratory, Oce@nyd, which is a part of a national project DIGITAL OCEAN for the distribution and creation of multimedia applications (audio, video, text etc.). In fact, it is designed to enable the public to discover, online, underwater environments. We can see in Figure 5 the Oce@nyd groupware and how it relates to other modules of the global project. In fact, the DIGITAL OCEAN project is constituted of three main sub-projects: the collaborative system (Oce@nyd), the 3D application (Oceanyd), and an underwater computerized display system conceived for swimming pools, associating video game, joysticks and multisensory simulation, called Tryton that is created by another partner of the project, as we can see in Figure 5. The aim of this global system, that is conceived and will be installed in regular swimming pools of any touristic organisation interested in it, is to enable the general public to visualise and navigate in real underwater sites reconstituted virtually by the 3D application Oceanyd that will be installed on the Tryton, and this with complete safety to the users that fear doing real scubadiving. In fact, the visualization – *in immersion*, of underwater contents, integrating virtual and mixed reality, breaks new ground in multimedia applications.

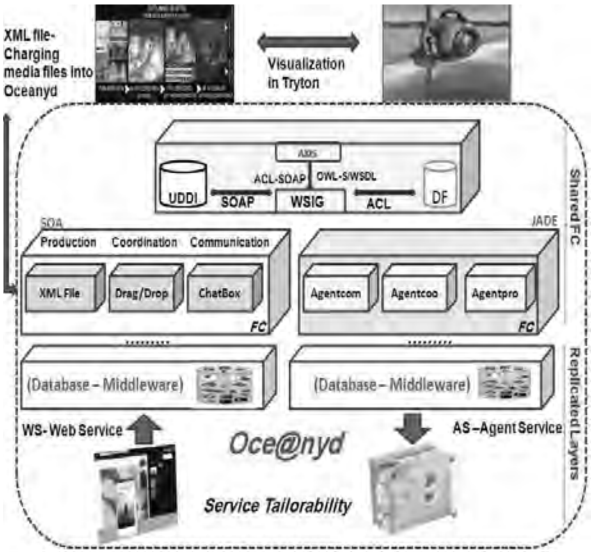


Figure 5: Implementation architecture of Oce@nyd

At the conceptual level we applied the FC shown in Figure 3 that includes two shared layers. At the implementation level, Ocea@nyd is a client\server application deployed on a Netbeans platform using JADE libraries, along with other libraries for implementing the Web services environment in the system. Both shared layers of the FC are deployed on the server side and other layers are replicated on users' machines, while the client/server communication is based on network streams. We will discuss mainly the FC of the model in a real application along with the physical interaction layer and the services offered that are tailored by users, while we make no assumptions, as we have mentioned before, for the other layers between the FC core and the physical interaction.

4.1 Web Interface

At the physical layer of Oce@nyd, lays a Web interface manipulated by users in collaboration. This web interface enables users to drag/drop multimedia files on a map of a specified underwater site taken by another partner of the project, using sensors, cameras and a GPS. The aim is to enrich this map with multimedia files by professional divers or users having real photos taken underwater when scuba diving in this particular location. We assume for now that the interface offers three mechanisms implemented with Web services technologies, each dedicated to one aspect of the 3C model as follows:

- **Communication**: the application provides a chat mechanism enabling users to exchange information about the files dragged onto the loaded map.
- **Coordination**: a service is encapsulated into the system that divides the map into variables zones, and which will detect the coordinates (x, y) of the dragged file (the number of zones dividing the map will be defined by the user himself, hence insuring tailorability of coordination capabilities in the system).
- **Production**: the system provides a mechanism for dumping the information of the multimedia applications dropped by users, into an XML file defining data attributes for every file dragged onto the map: date/time, nature (image, video or text), coordinates (x, y), description (user's description of a particular file). We will discuss in more details the information transmission via this XML generated file between Oce@nyd and Oceanyd in section 4.4.

4.2 Shared Server

These services discussed in the section above, reside in the functional core of the system, more particularly in the SOA environment residing on the server, while their WSDL files reside in the UDDI implemented one level up in the conceptual model, along with the DF and the set of codec provided by using JADE libraries. The main aim is to ensure consistency by binding one agent to its corresponding Web service via the WSIG as described in the previous section. For example, a communication Web service is managed by a communication agent insuring its integration and invoking in the system's physical interface. The part constituting the FC that contains all the agents are handled by the WSIG and JADE tools, providing basic methods for the registration of agents and their communication with the SOA. The agents use classes inherited from various behaviors offered by JADE libraries: CyclicBehaviour, Achieve REInitiator/ AchieveREResponder and Contract NetInitiator / ContractNetResponder.

4.3 Using functionalities of external Web services

In the case where the system is prompt for an external Web service, the user can specify the functionalities that he/she wants to use. Non-functional attributes can be specified such as (service quality, speed etc.) in order to select the Web service containing the functionality that best fit the user's needs. An agent handling the particular class of the Web service (communication, coordination or production) sends an ACL Request message to the WSIG containing the identity, name or any parameters identifying the Web service to be invoked. Received ACL messages are parsed and a SOAP message is constructed in order to prompt external public registries for this particular Web service, using the WSDL of the service to be invoked. If the Web service is found and a response is expected, a temporary endpoint is established on the Web server in order to receive responses, where the incoming SOAP message is parsed into an ACL Inform message and sent to the invoking agent, while registering the service in the local UDDI as a regular Web service. This particular agent will then be responsible for offering the requested Web service to the user as an agent service, while the process will be totally invisible for the user that is manipulating the Web interface.

Hence, on the design level, Web services can act as semi-autonomous agents that can be employed for describing the external behaviours and services discovered on the web and offered by software agents, and where agents build the system and constitute its mechanical behaviour, while their missions would be to search for available Web services and present them to users in collaboration as agent's services, based on users' preferences. In consequence, agents are used internally to establish high level, flexible and a dynamic interaction model, while the Web services will be more appropriate for resolving the problems of interoperability in the system, hence emerge the integration's synergy of the two worlds.

4.4 Link between the two applications

In order to situate Oce@nyd to the rest of the project, it is, as mentioned earlier, a collaborative platform that has two main aims: first, to present to users a new and innovative collaborative architecture that is specifically designed to support tailorability of services present in the system. By a service we mean any functionality that could enhance user's collaboration (chat, video stream, calendar, calculator, etc.), and the second aim is to collect multimedia files (images, videos, text, etc.) from users online in order to enrich 2D plans of underwater environments, for the essential purpose of charging these multimedia files into the offline 3D software Oceanyd, and thus keeping it up-to-date with new information. This data collection will charged into Oceanyd via an XML file containing all information relative to the multimedia files (coordinates, nature, time, etc.) collected by users. We can see a typical scenario in Figure 6 below:

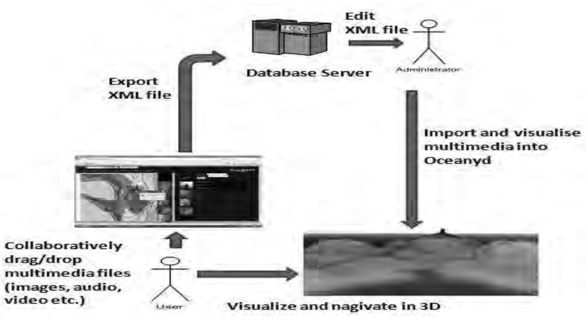


Figure 6: Typical scenario of information transmission between the two applications

A typical scenario starts with a user (we assume that a user is a professional diver or a diving club member that regularly does scuba diving) accessing his account in Oce@nyd in order to collaboratively enrich, using a drag/drop mechanism, 2D plan biocenose maps of underwater sites provided by a partner of the project, Semantic-TS (http://www.semantic-ts.fr). Upon connexion, the user will select from a list the underwater map that he/she wants to enrich with multimedia files (mostly images, video, audio and 3D objects). Before uploading the files, the user must fill out a form specifying the information concerning the files. More specifically, a user can specify the actual coordinates (specified by a GPS, or other underwater devices used by divers) of the location where the file is taken underwater, the time, temperature of the water, time etc. The user can also specify information about the site where the file was registered, such as, water turbidity, god rays, luminosity, etc. When finished, the user can then upload his/her file on the server, and drag it on a specific location on the map where the file was actually taken. When dropped, the file is given an (x,y) coordinates (unless the coordinates are already specified by the user), depending on the scale of the map. Of course, users enriching the map at the same time can collaborate in order to exchange information about the files being shared. For this, a chat mechanism is already implemented in the interface. Upon finishing the process of enriching the map, all information about the files dropped are saved, and an XML file is generated automatically that is put on the server in order to be edited by the administrator, while he can then check the files, delete irrelevant ones or modify others depending on their quality. The administrator then imports the XML file into the 3D Oceanyd application. The files are then put in their exact location in the 3D environment that corresponds to the 2D map enriched by users online. When the process is over, the user will then be able to use the Oceanyd application in order to visualize and navigate into the 3D underwater sites. A game, Nautilus Quest, is already developed in order for users to play together on a mission for finding, for example, files or 3D objects that they had contributed for to enrich a specific underwater site.

Of course, update scenarios must be studied in order to generate the XML file of all new information between two particular laps of time, and not recharging all the information every time in the same XML file, thus taking in consideration storage limit, query processing and performance time for charging the data into their exact location in Oceanyd, using the coordinates for every information dropped into the Web interface of Oce@nyd. The reason behind having two applications, one for the collect of information and the other for visualizing them, is that not all users have efficient bandwidth and internet speed in order to navigate online in 3D environments. Thus, dividing the process into, first, a collect of information using a 2D map, and then visualization on an offline 3D application constitutes a good solution to remedy this constraint.

In fact, these two sub-projects complete each other, while each one of them has a separate aim and purpose; their combination represents a coherent system to create a complete and realistic view of the underwater environment with real data taken undersea. The link between the two applications (Oce@nyd the online collaborative system, and Oceanyd the offline 3D application) is the information transmission generated by the former, and read by the latter as an XML file, in order to enrich and update the offline 3D environment with multimedia files collected by users online.

In what follows, we will talk about research concerns of the Oceanyd software, and that is terrain rendering to allow a more realistic view of underwater sites.

5. OCEANYD 3D VISUALISER

Real-time terrain rendering is a major preoccupation in 3D applications, as for example, virtual reality programs, video games, GIS etc. Indeed, considering the complexity of the problem, graphic cards' capabilities are still limited, despite the constant technological evolution. The main idea is to sensibly reduce the number of triangles in the scene, without sacrificing the initial geometry of the object. To remedy the problem, simple concepts have been already stated and discussed since the year 1976 by (Clarck, 76). In fact, an object situated far away from the point of view could be represented with less details (so with a smaller number of triangles) than a closer object. Following this initiative, a famous concept, Level-of-Detail (LOD), has been introduced. Moreover, each LOD of an object is loaded into memory and displayed according to his distance from the camera. This simple concept has to be adapted to terrain rendering. In this special case, the same object (the terrain) needs to be rendered with full details for the area near the virtual viewer and at the same time with fewer details for the rest of the terrain. Semantic-TS, one of the partners of the DIGITALOCEAN project, can produce bathymetric maps with biocenotic information, in supplement.

In fact, the model provided is a Digital Elevation Model (DEM) represented as a raster, which means on one hand, the points saved are regularly distributed on x and y axis, and on the other hand, each couple (x,y) is associated with an unique altitude z. The regular grid formed can thus be represented in many forms, while grayscale images of relief maps and triangulated geometric meshes of the surface being the two most employed types. In fact, many methods found in the literature have tried to solve the DEM rendering problem, we discuss the most important ones of our work in the following section.

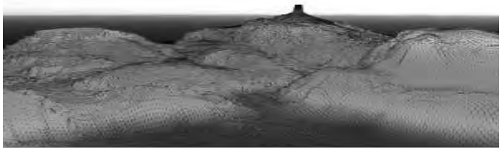


Figure 7: Oceanyd

5.1 Related Work

We can distinguish two steps in the evolution of the approaches. The first intention was to generate a mesh in real-time, following the needs, of all the terrain, respecting the previous constraints. Among the most famous algorithms, we find: Hoppe's approach (Hoppe, 96; Hoppe98), who proposed the Progressive Mesh, or the famous ROAM (Realtime Optimally Adapting Meshes) from Duchaineau (Duchaineau, 1997). A progressive mesh is created by a sequence of two operators: edge collapse and, its dual, vertex split. The sequence is in fact registered in order to be inversed later on. Moreover, Hoppe introduces a very interesting concept, the geomorphism, which removes all popping effect during LOD (Level-of-Detail) transitions. However, there is one major inconvenient: progressive mesh is relatively costly and is not adapted for complex scenes. On the other hand, The ROAM algorithm is much simpler. It uses a Binary Triangle Tree (BTT) structure to construct the mesh. Moreover, its simple implementation makes this method relatively popular and used often. The second generation of algorithms is based on, in addition to the definition of the different LODs, on the increasing performance of graphic cards and video memory that are integrated into. To restrict the cost of the choice in real-time, the level of visualization is determined for large triangulated areas than for each triangles. In this manner, (Losasso and Hoppe, 2004) have proposed the GeoClipmaps (inspired by the LOD definition of texture from (Tanner, 1998) the Clipmaps). In this approach, concentric strips at different level of detail are assembled and adapted around the point of view. Remedying the continuity between two strips that is being not ensured, the authors have proposed to make the failure invisible using a transparency fades on the geometry and the texture. In fact, this kind of approach is relatively new for terrain rendering. Prior to this work, complex operations were operated to avoid mesh's discontinuity.

5.2 A new problematic

After discussing with the divers community, one primordial element seemed missing in our prototype based on the DEM rendering: the excavations like overhangs, holes or caves. Indeed, the DEM doesn't permit to represent this kind of data since we should have several z for each couple (x,y). However, some solutions do exist:

– 3D objects are placed in a special position on the terrain to represent the needed excavations. In practice, the number of objects added limits the modelisation of underwater terrains.

– Another approach consists at representing some surface parts by some volumetric data. The meshes are generated with a classic algorithm as the Marching Cubes (Lorenson, 1987). As said previously, this algorithm is not adapted for underwater terrains with a lot of excavations.

In fact, our goal is to propose a terrain modelisation and visualization responding to all the constraints described earlier. We want to generalize the GeoClipmaps concept from Hoppe with the aim to adapt it to our model.

5.3 Model proposed and restrictions

Our suggestion is to describe a generalization of the initial DEM. In fact, we want to allow several altitudes (z) for every couple (x,y); however, the z values authorized will be distributed along a regular sampling. The 3D surface modeled, with excavations, is approximated by a group of three coordinates (x, y, z) regularly sampled. In addition, with the aim of keeping the surface's continuity, we propose to locally restrict the authorized number of configurations: with a regular paving of the space (for example, a 3*3*3 cube), we will choose some allowed situations, which will limit the ambiguities linked to the model's exploitation. This will define a LOD, noted *nk*. It is in fact with this notable restriction that we want to interactively edit the terrain.

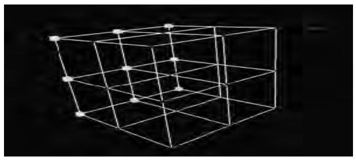


Figure 8: Example of a basic case- a wall

The next step of our work is to construct, from the *nk* LOD, the "next" LOD *nk+1*. Moreover, each cube from the *nk* LOD will give a new point (x, y, z). With this approach, all new positions will be adapted to match the authorized configurations, transposed to the new LOD. During the rendering, the different LOD exploitation will be done in function of the position and depending on the line of the viewer's site.

6. CONCLUSION

In this paper, we proposed a new preliminary architectural model that supports tailorability in the CSCW domain, where we apply it on an ongoing project Oce@nyd, which is itself a part of the bigger and global project, DIGITAL OCEAN. Our model relies on the integration of Web services and software agents that build system's components, and offers interoperability between heterogeneous applications by providing a synergy of technology used for the dynamic discovery and integration of Web services. This leads to conceiving a totally innovative approach, where the research field about web services and agent's integration is, until now, never been exploited in the context of groupware tailorability, and hence bringing innovation in the CSCW domain. Moreover, we have discussed research concerns for Oceanyd, and that is terrain rendering, where a new approach have been introduced to enhance visualization of 3D underwater sites in the application.

However, the XML file that is used to exchange the multimedia files collected by users and loaded into Oceanyd, might not be an efficient solution, as thousands of files might get uploaded on the server, and so managing all these information would be a quit hectic task. Therefore, other means of transmitting the information are to be discussed, as sharing the same database, and hence SQL queries might be optimised the search and load the files needed into Oceanyd.

In our future work, we aim to complete our implementation for the model. Our concern will be to modify incoming SOAP messages' headers of external Web services in order to fasten their integration into the Oce@nyd architecture according to their functionalities (communication, coordination and production). We believe that our preliminary approach for groupware tailorability will continue to mature through the use of Web services and software agents, which revealed to be appropriate to bring this concept from theory to practice.

6.1 References and/or Selected Bibliography

References from Books:

Newcomer, E., 2002. *Understanding Web Services: XML, WSDL, SOAP and UDDI*, Addison-Wesley Professional (E).

References from conference proceedings:

Bass. L., 1992 A metamodel for the runtime architecture of an interactive system. *User Interface Developers' Workshop*, *SIGCHI Bulletin*: 24(1).

Board, J., 2005 JADE Web Services Integration Gateway (WSIG) Guide*, Whitestein Technologies AG*.

Cheaib, N., Otmane, S., and Mallem, M., 2008. Integrating Internet Technologies in Designing a Tailorable Groupware Architecture. *12th IEEE Proc CSCWD*: pp 141-147.

Clark, J.H., 1976 Hierarchical geometric models for visible surface algorithms*. Communications of the ACM*, 19(10):547–554.

Dewan, P. 1998. Architectures for Collaborative Applications*. Computer-Supported Cooperative Work Journa*l, Trends *in Software, John Wiley & Sons*: 169-194.

Duchaineau, M., Wolinsky, M., Sigeti, D.E., Miller, M.C., Aldrich.,M, and M.B.,Weinstein 1997.. *Roaming terrain : Real-time optimally adapting meshes*. IEEE Visualization '97.

Ellis, C.A., Gibbs, S.J., Rein, G.L., 1991. Groupware: Some Issues and Experiences. C*ommunications of the ACM Journal.* ACM 34(1): 39-58 (1991).

Foukarakis I.E., Kostaridis A.I., Biniaris C.G., Kaklamani D.I and Venieris I.S., 2007. *Webmages*: An agent platform based on web services**.** *Computer Communications Journal*, Volume 30, Issue 3, pp. 538-545.

Gannod, G.C., Burge, J.E. and Urban, S.D., 2007. Issues in the Design of *Flexible and Dynamic Service-Oriented Systems. Proc* SDSOA'07: ICSE, IEEE. Washington (2007).

Hoppe, H., 1996. *Progressive meshes*. SIGGRAPH '96.

Hoppe, H., 1998. Smooth viewdependant refinement of progressive meshes*. Proc IEEE Visualization '98*, 1998.

Khezami, N., Otmane, S., and Mallem. M., 2005 A new Formal Model of Collaboration by Multi-Agent Systems. *Proc IEEE KIMAS*: 32-37, Massachusetts, USA.

Laurillau Y., and Nigay L., 2002. Clover Architecture for groupware, *Proc CSCW*, ACM Press (2002), pp. 236-245.

Lorenson, H.C, Marching, W., 1987. *cubes : A high resolution 3d surface construction algorithm*. Computer Graphics, 21.

Losasso, F., and Hoppe, H., 2004. *Geometry clipmaps : Terrain rendering using nested regular grids*. ACM SIGGRAPH.

Maamar, Z., Sheng Q.Z, and Benatallah B, 2003. *Interleaving Web Services Composition and Execution Using Software Agents and Delegation. AAMAS Workshop* (2003).

Maximilien, E.M and Singh, M.P., 2004. *A Framework and ontology for dynamic web services selection*. Internet Computing IEEE, Volume 8, Number 5 (2004), 84-93.

Morch, A., 1997. *Three levels of end-user tailoring: customization, integration, and extension*. Journal in Computers and design in context, MIT Press, 51-76.

Nguyen T.X., and Kowalczyk R., 2005. "*WS2JADE: Integrating Web Service with Jade Agents"* Technical Report, SOCAB05, Copyright Springer.

Peltz C., 2003. *Web Services Orchestration. A review of emerging technologies, tools and standards***.** Hewlett Packard White Paper.

Shen, W., Hao, Q., Wang, S., Li, Y., and Ghenniwa, H, 2007. Agent-based service-oriented integration architecture for collaborative intelligent manufacturing· *Robotics and Computer-integrated Manufacturing***,** Volume 23, Number 3, Elsevier (2007), 315-325.

Slagter, R., Biemans, M., and Hofte, H.T, 2001. Evolution in Use of Groupware: Facilitating Tailoring to the Extreme. *Proc. CRIWG:* 68-73.

Slagter, R., and Biemans, M., 2003. Designing Tailorable Groupware for the Healthcare Domain. *Proc. CRIWG, Springer* (2003), 58-73.

Stiemerling, O., and Cremers. A., 1998. Tailorable Component Architectures for CSCW- System**s.** *6th euromicro Workshop on Parallel and Distributed Programming*: pp. 21-24.

Tanner, C.C., Migdal, C.J., and Jones, M.T, 1998. The clipmap: A virtual mipmap. *SIGGRAPH '98*.

References from websites:

W3C, 2004. *Web ontology Reference Language.* http://www.w3.org/Submission/OWL-S/

Tilab. 2002. *JADE (Java Agent Development Framework),* Telecom Italia Lab http://sharon.cselt.it/projects/jade/

W3C, 1994. *World Wide Web Consurtium*, http://www.w3.org/

FIPA, 2008. http://fipa.org/

6.2 Acknowledgements

This project is supported by the National Agency of Research in France (ANR).

DEVELOPMENT AND EVALUATION OF ASUKA-KYO MR CONTENTS WITH FAST SHADING AND SHADOWING

Tetsuya Kakuta [a] *, Takeshi Oishi [a], Katsushi Ikeuchi [a]

[a] The University of Tokyo, Interfaculty Initiative in Information Studies, Graduate School of Interdisciplinary Information Studies, 7-3-1 Hongo Bunkyo-ku Tokyo, Japan - (kakuta, oishi, ki)@cvl.iis.u-tokyo.ac.jp

KEY WORDS: Mixed Reality, Augmented Reality, Photometric Consistency, Cultural Heritage, Subjective Evaluation

ABSTRACT:

We developed Mixed Reality (MR) contents that reconstructed the ancient capital of Asuka-Kyo and applied a fast shading and shadowing method that used shadowing planes. We conducted a subjective evaluation experiment with Head Mounted Display, which showed that displaying these contents increased the audience's knowledge of both Asuka-Kyo and MR technologies. We also conducted impression evaluation tests with and without shading and shadowing. These tests confirmed that shading and shadowing increased the audience's favorable responses to the evaluating factors of "Realistic," "Spectacular," and "Entertaining."

1. INTRODUCTION

Virtual Reality (VR) technologies enable us to digitalize cultural heritage objects as 3D reconstructed models (Utsugi et al., 2001; Tanikawa et al., 2004; Callieri et al., 2006). Using internet and VR systems, we can release to the public the digital contents of these objects that are normally confined to a museum or located in their original sites. These digital contents can then be utilized by educational, entertainment, and tourism industries.

Superimposing reconstructed 3D models of cultural heritage objects on their original sites using Mixed Reality (MR) has been widely carried out in recent years (Azuma, 2001). There are some restrictions on conventional VR contents. One limit is that we cannot appreciate them outside PC monitors and VR theaters. Another is that they lack a sense of reality compared with the original structures in their natural environments. On the other hand, MR allows us to make use of the landscape and atmosphere of the real environment and provide better experience for the users. Discovered ruins should be kept as close to the originals as possible for purposes of archaeological studies. MR can achieve effective representation of a reconstructed 3D model without having much effect on the original site and can thus strike a balance between preservation and presentation.

Many MR systems have been developed to virtually reconstruct 3D models of lost cultural heritage objects. Remains of ancient Greece have been reconstructed by the ARCHEOGUIDE project (Vlahakis et al., 2002). The LIFEPLUS project proposes 3D reconstruction of ancient frescos and paintings of Pompeii (Papagiannakis et al., 2002). Some old wooden Japanese buildings, the Nara palace site (Tenmoku et al., 2004), and the Kawaradera temple (Kakuta et al., 2004), are also represented by using MR technologies.

For the MR systems that aim to restore archaeological sites outside, it is important to achieve consistency of illumination. The shading of the virtual objects needs to match that of other objects in real environments. The virtual objects should also cast a correct shadow onto the real scene. Several methods have been proposed to solve this photometric problem in MR. Jacobs et al. provided a detailed survey of illumination techniques for MR (Jacobs et al., 2004). However, most methods are demonstrated on indoor scenes, and few of them are carried out at interactive update rates. Kakuta et al. suggested a novel method to express soft shadows of virtual objects in real time (Kakuta et al., 2007). This method is considered appropriate for static models in outdoor scenes, but it does not include actual applications for cultural heritage objects or subjective evaluation.

The evaluation of subjective psychological factors related to the illuminant consistency in MR is not emphasized in present research projects. Sato et al. reported an evaluation of shadow error between real and virtual occlusion objects (Sato et al., 2003). Sugano el al. also pointed out the importance of shadow representation for increasing object presence (Sugano et al. 2003). Photometric inconsistency on human faces was evaluated by Takemura et al. (Takemura et al., 2006). Though they performed a quantitative evaluation test for illuminant consistency, there was no discussion on what subjective psychological factors contributed to the sense of reality and presence, or how we could increase the overall satisfaction level of users with the contents. Furthermore, these cases were executed in an indoor environment, and subjective evaluation for actual MR contents in outdoor scenes was not carried out.

In this paper, we report an advanced evaluation test for illuminant consistency in MR. The motivation of this research is to identify psychological factors that influence evaluation of the contents and evaluate the subjective effectiveness of shadow representation using actual MR contents in an outdoor scene. We developed the MR contents of Japan's oldest capital, Asuka-Kyo, and applied the real-time illumination method (Kakuta et al., 2007) to this content.

Our reconstruction target, as stated, was the ancient capital of Asuka-Kyo, which is said to have existed in the Asuka district around the sixth and seventh centuries. The evaluation test was carried out as part of an open demonstration for the public on site, and the evaluators were selected from general tourists. Then we compared the learning effect before and after experiencing the MR contents and evaluated the improvement of their knowledge and interest in Asuka-Kyo and MR

(a) Current state of Asuka-Kiyomiharanomiya (b) Reconstruction image

Figure 1: Current photo of Asuka-Kiyomiharanomiya and its reconstruction image

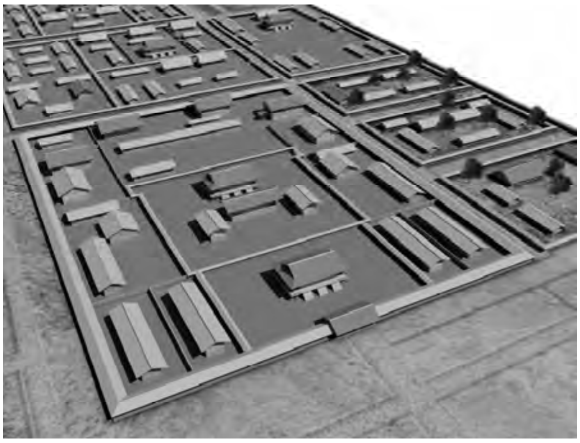


Figure 2: CG model of Asuka-Kiyomiharanomiya palace

technologies. The effectiveness of shadow was evaluated by the semantic differential method and factor analysis using synthesized images with and without shade and shadow on the MR system.

The rest of the paper is organized as follows, In Section 2, we introduce the outline of Asuka-Kyo and the process of constructing its MR contents. Then we describe the fast shading and shadowing method that we apply to the contents. In Section 3, the educational effect of the proposed contents is evaluated. The effectiveness of shading and shadowing for MR is discussed in Section 4. Finally, in Section 5, we present concluding remarks.

2. ASUKA-KYO MR CONTENTS

In this section, we present a brief overview of Asuka-Kyo and the process of creating MR contents. The fast shading and shadowing method that we apply to the contents is explained in the next section.

2.1 Reconstruction of CG model of Asuka-Kyo

Asuka-Kyo indicates the generic name of the capital where the palaces of various emperors were assumed to be in the Asuka area from the late sixth to the late seventh centuries. Though many tourists visit the historic sites of Asuka village, most of them cannot visualize the original sites and there is no way other than to rely on imagination to visualize the spectacle of Asuka at that time. Since Asuka village is a part of the application region of the "Law Concerning Special Measures for Preservation of Historic Natural Features in Ancient Cities"

Figure 3: CG model of Ebinoko great hall

and there is strict development control, it is difficult to build a new museum or replica for tourist. Moreover, excavation still continues and there is a possibility that the restoration plan will change with a newly-discovered archaeological fact. Against such a background, MR technology is very effective in reconstructing the ruins because it can overlay virtual buildings on the real site without harmful effect.

The restoration area in this project extends about 1350 meters in the north-south direction and about 600 meters in the east- west direction. Our MR contents contain the Asuka-Kiyomiharanomiya palace, the Kawaradera temple, the Asukadera temple, the Ishigami site, and the Asuka Enchi site.

It is known that there are multilayered remains in the palace area. We also decided to restore the Den-Asuka-Itabukinomiya III-B term (672-694) because it is in the upper layer and the position of the buildings is already known. The III-B term called Asuka-Kiyomiharanomiya differs from the III-A term called Nochino-Asuka-Okamotonomiya in the existence of the additional Ebinoko great hall in the southeast region. The Asuka-Kiyomiharanomiya palace expands 720 meters in the north-south direction and 100-450 meters in the east-west direction. It is said that there were government offices and gardens around the palace.

To create the CG model of Asuka-Kyo, we referred to the reconstruction drawings provided from the Nara National Research Institute for Cultural Properties and the Archaeological Institute of Kashihara. We modelled CG buildings of Asuka-Kyo with 3dsmax by Autodesk Inc. manually. The painted surface, tiled roofs of temples and the plain wood material of palace buildings are expressed by texture images. Figure 2 shows the reconstruction model of the Asuka-Kiyomiharanomiya palace, and Figure 3 shows that of the Ebinoko great hall. Figure 1 shows the synthesized image of the Asuka-Kiyomiharanomiya palace and the current landscape of the Asuka village. The total number of polygons in these models is about 686,000.

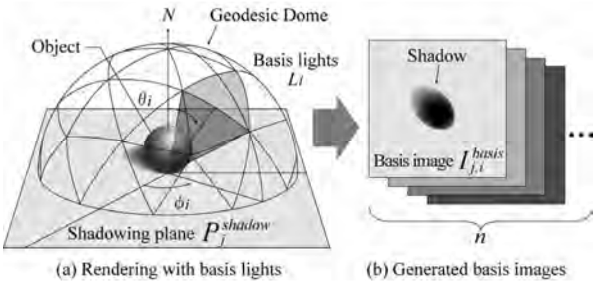


Figure 4: Generation of basis images

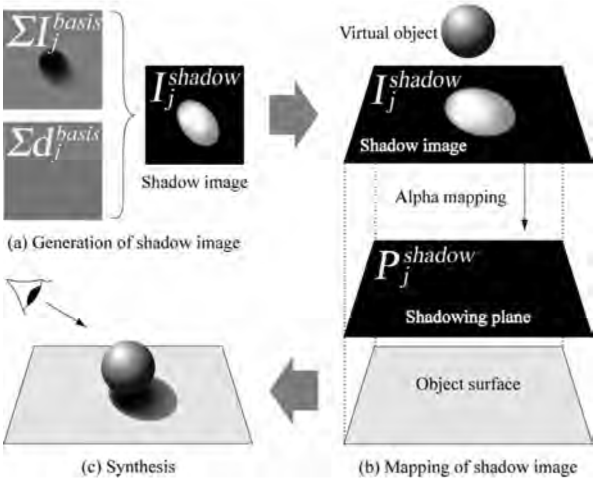


Figure 5: Generation of shadow images and mapping onto shadowing planes

The requirements for the Asuka-Kyo MR content are:

1. Rendering large-scale complex models in real time.
2. Realistic rendering for a close look by users.
3. Realistic illumination corresponding to the change of illumination outside.

However, there is a general trade-off between quality and efficiency in the rendering process. Conventional methods (Haller et al., 2003; Jacobs et al., 2005) can express only hard shadows in fixed illumination. We applied the fast shadowing method to the proposed contents in order to fulfil these three requirements. We found it is easy to apply the shadowing plane method proposed in (Kakuta et al., 2007) to the models of Asuka-Kyo because they are basically static and have a simple form without complex wooden parts. Also they consist of only diffuse materials without specular components, so it is possible to generate correct shadows from basis images.

2.2 Shading and Shadowing

The shadowing method that we applied to the proposed MR contents can express soft shadows of virtual objects efficiently. The method renders basis images offline and synthesizes them to create soft shadow images corresponding to the illumination of the real environment. We describe this shadowing process briefly.

2.2.1 Generation of Basis Images: In order to obtain the basis images, we set up the shadowing planes $P_j^{shadow}\{j=1,2,\ldots,m\}$ on the geometry of the scene. Shadowing planes are generated from convex hulls of each virtual object. Each shadowing plane covers virtual objects roughly, and is offset a little in the direction of a user's viewpoint to avoid

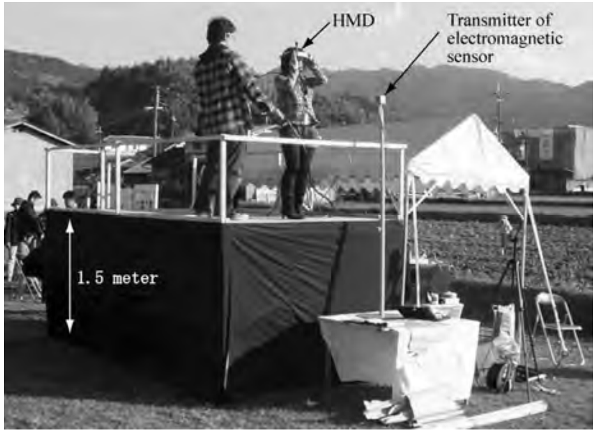


Figure 6: Appearance of evaluation experiment

z-fighting error. Then we assume the illumination of the scene is a hemispheric surface light source. We divide this surface light source by the geodesic dome model and generate basis lights $L_i\{i=1,2,\ldots,n\}$ as shown in Figure 4. Using each basis light, we render the scene and obtain basis images $I_{j,i}^{basis}$ and diffuse components $d_{j,i}^{basis}$ on the surface. The basis images show the shadow of objects in the scene.

2.2.2 Mapping to Shadowing Planes: We obtain the illumination of the real environment from a video camera with a fisheye lens. The radiance parameters that indicate the distribution of scene radiance are computed from an omnidirectional image taken by the camera. Using this radiance parameter, we control the intensity of scene light and are able to express correct shading of virtual objects. Next we compute the linear combination of the basis images and radiance parameters and generate a soft shadow image that corresponds to the scene illumination. Finally, generated soft shadow images are mapped onto previously-installed shadowing planes as an alpha texture. Rendering the scene with these shadowing planes, we can express soft shadow images of virtual objects (Figure 5). Note that we can compute the linear combination of basis images on a standard Graphical Processing Unit (GPU). As a result, the soft shadow images are generated very fast, and this method can be implemented for a real-time interactive MR application.

2.3 Experimental Setup

2.3.1 MR system: Our MR system is mainly based on Canon's MR Platform system (Uchiyama et al., 2002), which includes a video see-through head mounted display (HMD) and the Polhemus's Fastrak, six degree-of-freedom (DOF) electromagnetic tracking sensor. This MR system enables us to acquire the image of the real site, align virtual objects to the image, and synthesize them. The size of the image displayed on HMD is 640x480 pixels. The evaluation tests to be described are performed on a standard computer system with an Intel Core 2 Duo E6850 CPU, NVIDIA GeForce 8800GTS GPU and 4GB RAM. For obtaining the illumination, we use Victor's CCD video camera KY-57 equipped with a fisheye lens (Fit Corporation's FI-19).

2.3.2 Displaying Procedure: We carried out subjective evaluation tests on the historical site of Asuka village. Since the reconstructed CG model of Asuka-Kiyomiharanomiya contains wooden fences 2 meters high in the inner bailey, we put up a scaffolding 1.5 meters high so the user could overlook the

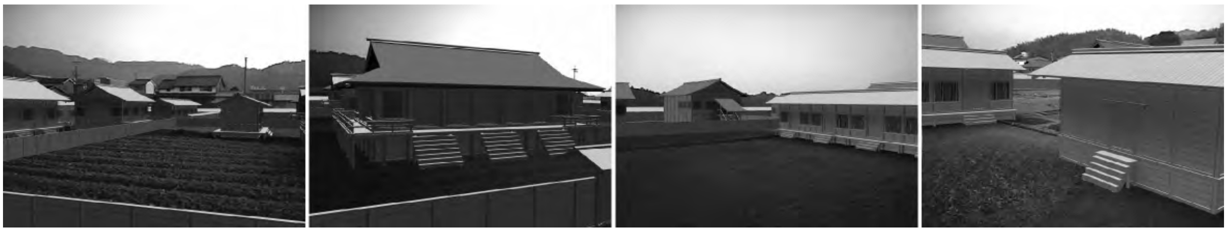


Figure 7: Synthesized image of Asuka-Kiyomiharanomiya palace on MR system

End points	Question
1. Raising interest in Asuka-Kyo and local community.	Q1. Do you want to know the history and the culture in the Asuka Period? Q2. Do you want to visit other archaeological sites in Asuka village?
2. Acquisition of new knowledge of the history and culture of the Asuka Period.	Q3. Do you feel as if the ancient Asuka-Kyo were here? Q4. Do you know what the buildings of the Asuka Period were? Q5. Can you imagine people's livelihood in the Asuka Period? Q6. Can you imagine how large the Asuka-Kyo was?
3. Getting the importance of the preservation of cultural properties.	Q7. Do you think that the preservation of cultural heritage is important? Q8. Do you think the reconstruction using Mixed Reality is meaningful?
4. Getting the mechanism and property of CG and MR.	Q9. Do you understand the relation between display and viewpoint? Q10. Do you understand the mechanism of Mixed Reality?

Table 1: Relation of end points of education effect and questions

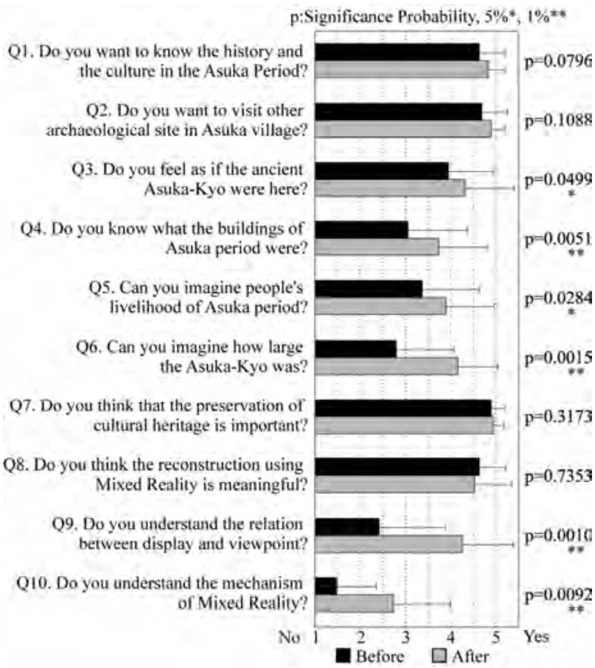


Figure 8: Average score of educational effects before and after experiencing the contents

the palace as shown in Figure 6. We installed the MR system on a northeast direction of the main temple because there are no real buildings and the users could take a look around the site. The camera for illumination acquisition was placed 5 meters away from the MR system to avoid capturing users and equipment. Figure 7 shows the synthesized images on HMD. Natural soft shadows and real-time image synthesis were achieved by application of the fast shadowing method.

3. EVALUATION EXPERIMENT FOR EDUCATIONAL EFFECT

First we performed an experiment to evaluate the educational effect of the proposed contents by comparing the attitudes of subjects before and after experiencing the contents. We wanted to investigate how the test subjects were affected by the contents in respect to their interest in and concerns about the history of Asuka Period.

3.1 Experimental Design

Nineteen observers (12 male, 7 female, average age 49.63, range 28-77) volunteered to participate in the experiment. Evaluation items were determined from the learning goals of Geography, History, and Information subjects in the ministry's curriculum guidelines for high schools. Specifically, we presumed four end points of 1) Interests in the remains, 2) Historical knowledge, 3) Conservation awareness of cultural heritage objects, and 4) Understanding of MR technologies, which were expected to increase by experiencing the MR contents. Ten questions were prepared from these end points. The corresponding relation of questions and end points is shown in Table 1. These questions were randomly arranged on the survey forms, and observers were asked to rate them using a scale from one to five before and after experiencing the contents. The experience time of each observer was two minutes.

3.2 Result and Discussion

As a result, we confirmed the score had been increased for almost all question areas. Figure 8 shows the average score of every question before and after going through the contents. The p-values in Figure 8 indicate the significance probability calculated by the Wilcoxon signed-rank test, and they represent the probability of obtaining the score by chance (two-sided test). The symbols of * and ** indicate a significance level less than 5% ($p<0.05$) and 1% ($p<0.01$) respectively. We found a statistically significant improvement of less than 1% in Q4, Q6, Q9, and Q10, and less than 5% in Q3 and Q5.

The result shows the proposed contents had an educational effect on learning goals 2) Historical knowledge, and 4) Understanding of MR technologies, previously assumed. Q3 and Q6 showed how users increased their recognition of the scale of ancient Asuka-Kyo and its spatial corresponding relationship with current remains and landscape. Especially the score of Q6 regarding size is remarkably increased and we

(a) Synthesized image without shading and shadowing (b) Synthesized image with shading and shadowing

Figure 10: Synthesized images for evaluation experiment for user's impression with and without shading and shadowing

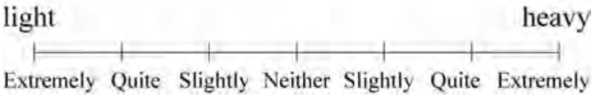


Figure 9: Sample of semantic differential scale

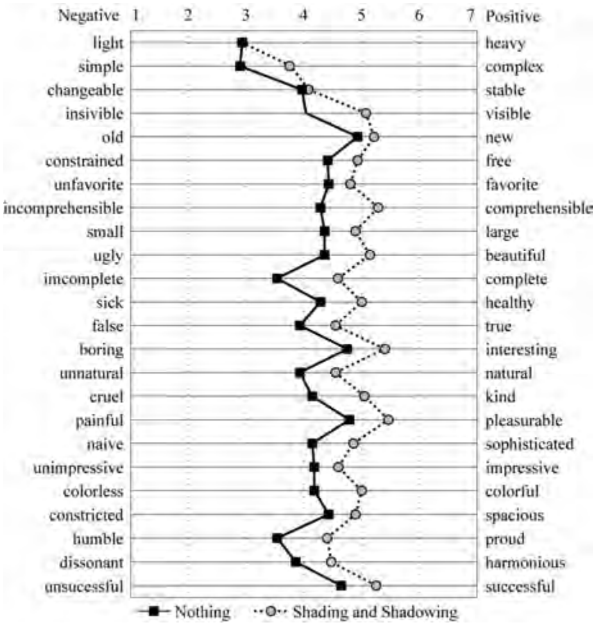


Figure 11: Average score of adjective pairs

Adjective pairs	Factor loading		
	Factor 1	Factor 2	Factor 3
Factor 1: "Realistic"			
false – true	**0.947**	0.075	-0.171
incomplete – complete	**0.677**	-0.060	0.316
dissonant – harmonious	**0.628**	0.052	0.095
humble – proud	**0.619**	0.079	0.287
unnatural – natural	**0.546**	0.134	0.287
sick – healthy	**0.535**	0.204	0.206
Factor 2: "Spectaculer"			
colorless – colorful	0.057	**0.840**	0.034
unimpressive – impressive	0.187	**0.815**	-0.064
unsuccessful - successful	-0.108	**0.706**	0.228
naïve – sophisticated	0.489	**0.551**	-0.090
constricted – spacious	0.190	**0.517**	0.041
Factor 3: "Entertaining"			
incomprehensible – comprehensible	-0.096	0.175	**0.868**
constrained – free	0.140	-0.093	**0.629**
unfavorite - favorite	0.065	0.236	**0.545**

Table 2: Factor loadings of adjective pairs

might guess it is because the observers could physically experience the spaciousness of Asuka-Kyo through the MR system. The scores of Q9 and Q10 show that the contents contributed to users' understanding of the MR technologies and devices. The average scores of these questions were 2.42 and 1.47 respectively. These scores were relatively low compared to other questions, but they dramatically increased after experiencing the MR system. Getting hands-on experience should be recognized as a means of increasing understanding of MR technologies.

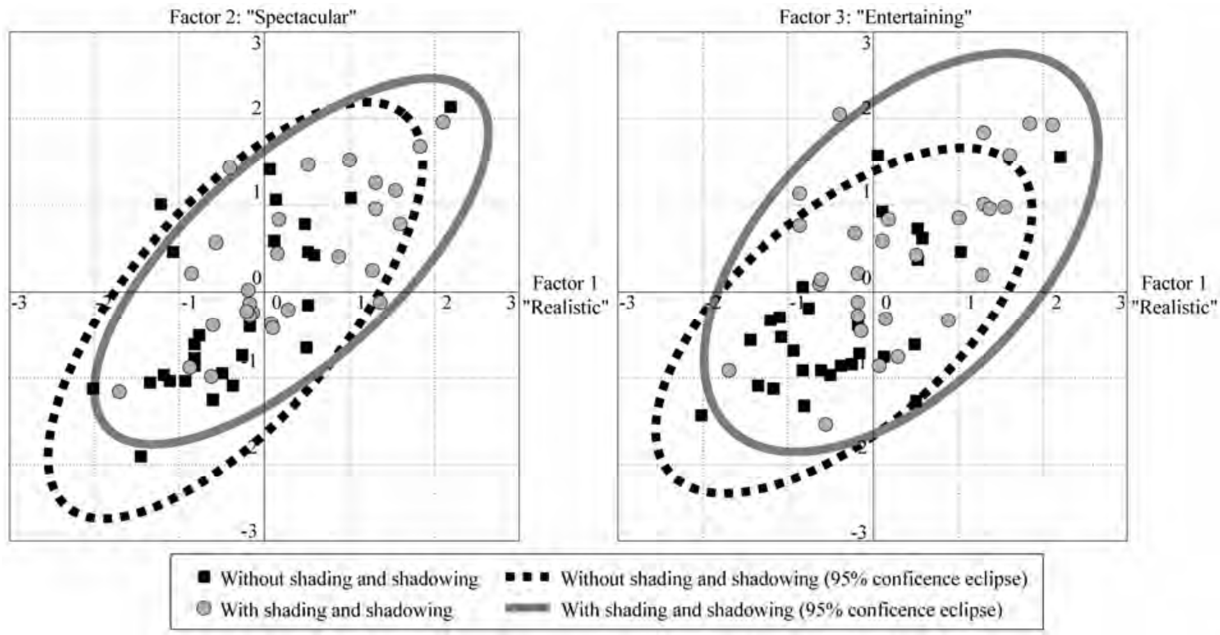
On the other hand, the scores of Q1, Q2, Q7, and Q8 showed no significant educational effect. We might assume this is because the interest and concerns of examinees were high enough from the beginning so that the average scores for these questions were positive and the contents had a very small effect on the observer. Especially, the average score of Q7, preservation of cultural heritage, was originally higher than that for other questions. Meanwhile the score of Q8, the value of mixed reality, decreased, but this might be because of random error (p=0.7353).

4. EVALUATION EXPERIMENT FOR USERS' IMPRESSION WITH AND WITHOUT SHADING AND SHADOWING

In the next experiment, we conducted a subjective evaluation test to measure the users' impressions of the MR contents and clear up the potential psychological factors that affected evaluation of the synthesized image. Then we studied whether significant differences came about by adding correct shade and shadow.

4.1 Experimental Design

Fifty-six observers (28 male, 28 female, average age 51.88, range 12-81) who were not engaged in the experiment of Section 3 volunteered to participate to this experiment.



(a) Factor scores in the axes of factors 1 and 2 (b) Factor scores in the axes of factors 1 and 3

Figure 12: Factor score of examinees with and without shading and shadowing

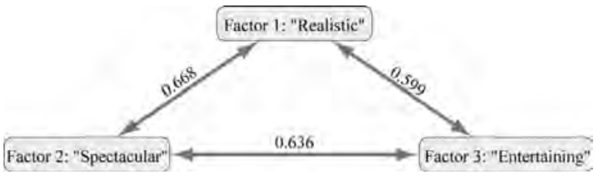


Figure 13: Correlation coefficient among factors

We asked them to report the impression of the synthesized images on the MR system. The semantic differential method was used to analyze the result, and we began with 24 pairs of adjectives that were considered to adequately describe the impression of a synthesized image based on (Osgood et al., 1967). The examinees were asked to rate the images using a scale from one to six. The example of the assessment measure is shown in Figure 9. The observers were divided into two groups and were shown two types of synthesized images on the MR system with and without shading and shadowing for each group (Figure 10). Note that we did not inform each group about the difference. The experience was performed under a clear sky within an hour to minimize the change of illumination condition. The experience time of each group was two minutes, like the experiment in Section 3.

4.2 Result and Discussion

Figure 11 shows the average score of each adjective pair with and without shade and shadow. The left side of the figure contains negative adjectives, and the right side contains positive adjectives. When we compared each average score, the group "with shading and shadowing" clearly moved to the right (positive) side. In particular, there was significant difference over 1.0 in the adjective pairs, "invisible – visible" and "incomplete – complete." This could be due to the fact that the shade and shadow emphasized the appearance of solidity and the models seemed to be authentic. By contrast, there was little difference in "light - heavy" and "changeable – stable." It would appear that the shade and shadow had an insignificant effect on the model's projection of stability.

The factor analysis reveals the potential common factors related to the contents evaluation as shown in Table 2. In the calculation process, we extracted the initial factors by the principal factor method and adopted a promax rotation. The adjective pairs in Table 2 are sorted by their absolute value of factor loadings. The result shows that the three extracted factors explain almost all the adjective pairs.

In the first factor, the factor loading of "false – true" pair is especially high (0.947) and it contains "dissonant – harmonious" and "unnatural - natural" pairs. Therefore, we assume this factor is related to the harmony and consistency of the image and it can be said to represent the "Realistic" aspect. In the second factor, the factor loadings of "colorless – colorful" and "unimpressive – impressive" are high; we assume this means the appeal and power that is "Spectacular." In the third factor, the factor loadings of "incomprehensible – comprehensible", "constrained - free" and "unfavorite - favorite" are high, so we consider this means an acceptance and appeal of the contents, in other words "Entertaining."

The factor scores of three factors described above for each group are plotted in Figure 12. The ellipses in this figure mean the 95% confidence region given a 2-dimension normal distribution of factor scores. When we compare the distribution of each group, "with shading and shadowing" group has a higher score than the other. Particularly, there is a great difference between two groups in the axis of the "Entertaining" factor as shown in Figure 12 (b).

Figure 13 shows the coefficient of correlation of the three factors. There are fairly high correlations among these factors, so we can consider the correct shading and shadowing in MR are important to improve not only the reality and consistency of the synthesized image but also the sharpness and punch of models, as well as the entertainment value and interactivity of the contents.

5. CONCLUSION

In this paper, we developed and evaluated MR contents of the Japanese ancient capital Asuka-Kyo. We successfully improved the reality of the synthesized image in this content by applying the fast shading and shadowing method (Kakuta et al. 2007). Our first subjective evaluation experiment shows the educational effect of the contents. MR applications that reconstruct lost remains on historical sites are shown to increase the interests and concerns of the subjects. The result of the second experiment regarding users' impressions reveals three common factors in evaluating MR contents, "Realistic," "Spectacular," and "Entertaining." We confirmed that the shading and shadowing of virtual objects improved these evaluating factors and the illuminant consistency is essential for the subjective evaluation. Correct shading and shadowing are also important to increase the entertainment value of the contents in addition to the reality and consistency. Improving the educational effect of the Asuka-Kyo MR contents is our task for the future. Especially the end points of "Historical knowledge" and "Conservation awareness for cultural heritage" are minimally affected by the MR contents as described in Section 3. Multiple information service such as text and audio commentary might improve these aspects.

6. REFERENCE

Azuma R., Baillot Y., Behringer R., Feiner S., Julier S., and MacIntyre B., 2001. Recent advances in augmented reality, *IEEE Computer Graphics and Applications*, Vol. 21, No. 6, pp. 34-47.

Callieri M., Debevec P., Pair J., and Scopigno R., 2006. A realtime immersive application with realistic lighting: the Parthenon, *Computer & Graphics*, Vol. 30, No. 3, pp. 368-376.

Haller M., Drab S., Hartmann W., 2003. A real-time shadow approach for an augmented reality application using shadow volumes, *Proc. Symp. on ACM Virtual Reality Software and Technology (VRST 2003)*, pp. 56-65.

Jacobs K. and Loscos C., 2004. Classification of illumination methods for mixed reality, *Proc. Eurographics State of the Art Report (STAR)*, pp. 95-118.

Jacobs K., Nahmias J. D., Angus C., Reche A., Loscos C., Steed A., 2005. Automatic generation of consistent shadows for augmented reality, *Proc. Graphics Interface 2005*, pp. 113-120.

Kakuta T., Oishi T, and Ikeuchi K., 2004. Virtual kawaradera: fast shadow texture for augmented reality., *Proc. Int. Society on Virtual Systems and Multimedia (VSMM 2004)*, pp. 141-150.

Kakuta T., Oishi T., and Ikeuchi K., 2007. Real-time soft shadows in mixed reality using shadowing planes, *Proc. IAPR Conference on Machine Vision Application (MVA 2007)*, pp. 195-198.

Osgood C. E., Suci G. J., Tannenbaum P., 1967. *The measurement of meaning*, University of Illinois Press, Urbana and Chicago.

Papagiannakis, G., Ponder, M., Molet, T., Kshirsagar, S., Cordier, F., Magnenat-Thalmann, N., and Thalmann, D., 2002. LIFEPLUS: Revival of life in ancient Pompeii, *Proc. Int. Conf. on Virtual Systems and Multimedia (VSMM 2002)*, pp.19-26.

Sato I., Sato Y., and Ikeuchi K. 2003. Illumination from shadows, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 3, pp. 290-300.

Sugano N., Kato H., and Tachibana K. 2003. The effects of shadow representation of virtual objects in augmented reality, *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR 2003)*, pp.76-83.

Takemura M., Kitahara I., Ohta Y., 2006. Photometric Inconsistency on a Mixed-Reality Face, *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR 2006)*, pp.129-138.

Tanikawa T., Ando M., Wang Y., Yoshida K., Yamashita J., Kuzuoka H., and Hirose M., 2004. A case study of museum exhibition: historical learning in Copan ruins of Mayan civilization-, *Proc. IEEE Virtual Reality Conf. (VR 2004)*, pp.257-258.

Tenmoku R., Nakazato Y., Anabuk A., Kanbara M., and Yokoya N., 2004. Nara palace site navigator: device-independent human navigation using a networked shared database, *Proc. Int. Conf. on Virtual Systems and Multimedia (VSMM 2004*), pp. 1234-1242.

Uchiyama S., Takemoto K., Satoh K., Yamamoto H., and Tamura H., 2002. MR Platform: a basic body on which mixed reality applications are built, *Proc. Int. Symp. on Mixed and Augmented Reality (ISMAR 2002)*, pp. 246-253.

Utsugi K., Moriya T., and Takeda H., 2001. Making of a virtual world of heijokyo from historical knowledge, *Proc. Int. Conf. on Virtual Systems and Multimedia (VSMM 2001*), pp. 455-462.

Vlahakis V., Ioannidis N., Karigiannis J., Tsotros M., Gounaris M., Stricker D., Gleue T., Daehne P., and Almeida L., 2002. Archeoguide: an augmented reality guide for archaeological sites, *IEEE Computer Graphics and Applications*, Vol. 22, No. 5, pp. 52-60.

7. ACKNOWLEDGEMENT

This research was, in part, supported by the Ministry of Education, Culture, Sports, Science and Technology, under the program, "Development of High Fidelity Digitization Software for Large-Scale and Intangible Cultural Assets." The experiment in Asuka Village was supported by the Ministry of Land, Infrastructure and Transport. We are also grateful to the village office of Asuka, the Archaeological Institute of Kashihara and the National Research Institute for Cultural Properties.

ACKNOWLEDGEMENT

This research was, in part, supported by the Ministry of Education, Culture, Sports, Science and Technology, under the program, "Development of High Fidelity Digitization Software for Large-Scale and Intangible Cultural Assets." The experiment in Asuka Village was supported by the Ministry of Land, Infrastructure and Transport. We are also grateful to the village office of Asuka, the Archaeological Institute of Kashihara and the National Research Institute for Cultural Properties.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

CYBER-ARCHAEOLOGY: AN ECO-APPROACH TO THE VIRTUAL RECONSTRUCTION OF THE PAST

M. Forte [a]

[a] School of Social Sciences, Humanities and Arts, University of California, Merced, P.O. Box 2039, Merced, CA 95344, US, mforte@ucmerced.edu

KEY WORDS: Cyber-Archaeology, Virtual Museum, Ecology, Communication, Villa of Livia

ABSTRACT:

This paper aims at introducing and discussing an epistemological model of cyber archaeology in relation to the need to investigate what happens in a immersive environment of virtual archaeology where every user is “embodied” in the cyber space. In particular the ontology of archaeological information, or the cybernetics of archaeology, refers to all the interconnective relationships which the *datum* produces, the code of transmission, and its transmittability. Because it depends on interrelationships, by its very nature information cannot be neutral with respect to how it is processed and perceived. It follows that the process of knowledge and communication have to be unified and represented by a single vector. 3D information is regarded as the core of the knowledge process, because it creates feedback, then cybernetic difference, among the interactor, the scientist and the ecosystem. It is argued that Virtual Reality (both offline and online) represents a possible ecosystem, which is able to host top-down and bottom-up processes of knowledge and communication. In these terms, the past is generated and coded by “a simulation process”. Thus, from the first phases of data acquisition in the field, the technical methodologies and technologies that we use, influence in a decisive way all the subsequent phases of interpretation and communication. In the light of these considerations, what is the relationship between information and representation? How much information does a digital model contain? What sorts of and how many ontologies ought to be chosen to permit an acceptable transmittability? These and many other questions on related topics take on certain urgency because they relate directly to the loss of information from understanding, learning, and the transmittability of culture. Indeed, our ability to transmit culture depends on a model which combines on the same axis processes of understanding and communication. Thus, the questions which we pose in a phase of bottom-up knowledge (for example, in an archaeological excavation) will influence the top-down phases of interpretation, or the mental patterns (for example, a comparative analysis and reconstruction of models). From this derives the need to interconnect the top-down processes with the bottom-up in accordance with a reciprocal systemic interaction, for example in a virtual space where both sequences can coexist. If we peremptorily separate knowledge and communication, we risk losing information along the way, reducing the relationships that are constructed between acquisition/input and transmission/output. Archaeological communication ought to be understood as a process of validation of the entire cognitive process of understanding and not as a simple addendum to research, or as a dispensable compendium of data. This theoretical discussion will be applied to a presentation of an important case study: the “Virtual Museum of the Ancient Via Flaminia” in Rome, a MUD space where diverse users can interact in the same cyberspace.

1. INTRODUCTION

1.1 The case of the Villa of Livia and the Virtual Museum

In January 2008, in Rome, the virtual multiuser Museum of Ancient Via Flaminia was open at the National Roman Museum (Museo Nazionale delle Terme di Diocleziano). One of the key archaeological sites reconstructed at different levels in the system is the famous Villa of Livia, wife of the emperor Augusts, located in the North-East part of Rome (Forte et alii, 2006, Forte, 2008). The virtual reality system consists in a virtual room provided with four interactive platforms. Users explore and share the virtual space through avatars: with their actions they create a virtual "show" which can be seen from the audience on a central screen (fig.1). There are two levels of exploration:

1. A free navigation where users can move into the space following their own personal curiosity, enterprise or perception;
2. A narrative level where audiovisual, digital reconstructions and narrative formats are displayed.

Figure 1. Graphic project of the Virtual Museum of the Ancient Via Flaminia (Roman National Museum).

On the main screen, through a general "direction", different visual and informative contents show what happens in the virtual environment through the movements of the users/avatars. The whole audio system is shared and has 3D stereophonic sounds. The "Virtual Museum of the ancient via Flaminia" and particularly the reconstruction of the Villa of Livia Drusilla is the first archaeological project developed through several media and technologies at the same time. The project's final aims are:

1. The reconstruction of a very important archaeological area, even if nowadays its consumption is very limited to the public: although open to the public, the villa is located out of the traditional touristic routes and totally covered by a permanent roof which prevents the public to have a clear and complete overview of it. The villa is therefore difficult to understand either its archaeological structure and as in its historical and cultural value.

2. The recontextualization of landscapes, objects and monuments concerning the Ancient "Via Flaminia" and the Roman National Museum

3. The virtual reconstruction which aims at effectively communicating complex data throughout a direct and detailed view of the entire area in different archaeological phases.

Figure 2: The archaeological site of the Villa of Livia (Rome). The coverage thwarts the comprehension of the monument.

1.2 Overview

The archaeology of the third millennium is able to process, interpret and communicate much more data and information than in the last two centuries. Are we aware of how much data can be produced and disseminated in this era? And how much fast is this process? In the 90's most part of research projects in virtual archaeology were technologically oriented; now we think that in the third millennium they should be cybernetic-oriented. These informative-cybernetic models represent the focus of the methodology of validation and the scientific-cultural content we would like to send to the future.

But we have to pay attention: most part of what we study is relativized from subjective interpretation, then discretized from the output restrictions (ex. paper's space, color and resolution of photos, limits of drawing, accuracy of data, etc.) but not fully perceived in a path of final validation. The capacity to transmit knowledge and interpretation depends on a complexity of diverse factors: format, accuracy, argument, induction-deduction, communication, context, ontology. The object of knowledge transmission is what is perceived processed and finally communicated according to a constructivist logic (Watzlawick, 1985). To Piaget the organization is always the result of a necessary interaction between conscious intelligence and the environment (Piaget, 1980).

A path of research archaeological exclusively of taxonomic type (mostly bottom-up) is not ever complete, since it is not aimed to the comprehension and communication of the context, while, on the contrary, a syntagmatic approach, based on a chain of codes, meanings and relations, is strongly perceivable and reconnect the original code with the context (Antinucci, 2004). In the field of the ecological thinking it can be explained as a relationship map-territory (Bateson, 1972, 1979) , where the map represents the information code (Korzybski, 1941) and the territory the information not yet coded. For example the archaeological landscape is a territory, while the ancient landscape is a map (Forte, 2003, 2005). Every archaeological context had from its origin a strong information-communication *autopoietic* content, that is able to produce meanings in its society, since the message is easily understood in its original context. Because of the spatial-temporal decontextualization, major part of this autopoietic code was lost; this is due to the missing meanings (for the observer and consumer) of the cultural and natural landscape and all the artificial relations with monuments and human actions. The archaeological research has from a long time enormous difficulties to face the scientific validation processes of the *datum*, mainly restricted to excavation reports or written publications. In terms of scientific validation how much is it possible to reconstruct of the long process of archaeological interpretation? Is the quality and quantity of produced information sufficient to validate the entire scientific pipeline?

The final response is in the validation system of data, in the transparency of interpretation processes, visualization and interaction and in the ability to codify and transmit information. In these terms we can consider Virtual Archaeology a cybernetic process and not a technological outcome. Our work was inspired by the second cybernetics of Gregory Bateson (1904-1980), by the study of relations between information, environment, organisms, ecosystem (Bateson, 1967), anthropology and ecology of culture (Ingold, 2000), by the concept of *affordance*-relation of the Gibson's thinking (Gibson, 1999). Following these premises we have studied the relationship between system and context of archaeological information. In particular we pay more attention to the cybernetic model (who follows rules of information transmission) than to the computer model. The cybernetic model makes informative models, while the computer model develops mainly tools of data processing. Therefore the cybernetic model represents a simulation process, an open virtual connective space where the information is generated by feedback's relations and by interaction.

In this paper I use the term "cyber-archaeology", preferred to "virtual archaeology" since I need to explain the complex ecological process/feedback used in the interaction with a virtual environment. Then I use the term "embodiment" to indicate the properties of interaction in a multiuser immersive virtual environment (Biocca, 1997; Gallese, 2005).

The creation of the "Virtual Museum of Ancient Via Flaminia in Rome, open in January 2008 (Forte, 2008), constitutes a good premise for discussing the role of cyber-(virtual)-archaeology in this digital age.

2. CYBER ARCHAEOLOGY

2.0 Virtual Archaeology

In the last decade the concept of Virtual Archaeology was discussed and popularized (Forte, 1997) with the description of many different scenarios. Most part of the discussions was focused on the value and potentiality of the digital reconstruction but I think that not enough attention was paid to the potentiality of the behavioral simulation processes. It is quite easy to follow the reconstruction process in an activity of digital modeling, but it is very difficult to explore the mental abilities of interaction in the cyperspace: here any action and feedback can produces new models of knowledge and interpretative processes. This condition of simulated and increased reality can be defined “hype-reality” (Baudrillard, 1994): for Baudrillard this kind of simulation is “more real than real”. Even if the Baudrillard’s interpretation of hyper-reality is very negative (the danger is that the Virtual can cancel the Real), this vision can help us to understand that the virtual represents a “dense”, augmented information.
According to Maturana and Varela the specific dynamics of interaction and embodiment are able to increase the capacities of learning. The feeling of immersion in the virtual world is generated by a multisensorial involvement and by the inclusion of the user in the 3D space (Richardson, Montello, Hegarty, 1999).
It is through the mind-body that it is possible to know the virtual world and, to a lower level, models and information are processed. The virtual reality systems, as cognitive technology, interpret successfully an enactive approach to the cognition, such as computer and artificial intelligence interpret the cognitive hypothesis ”(Morganti, Riva, 2006). The enactive cognitivism discusses the dichotomy between intern and extern: therefore cognition is an action “embodied” (Varela *et alii* 1991). In terms of enaction, the cognition depends on perceptual-motor experience and these capacities belong to a wider biological psychological and cultural context.
Thus the issue of the information’s acquisition would be identified in the circularity between action and experience and between action and knowledge (Varela *et alii* 1991). Every existing object in the world depends on this perceptual-motor interaction. The object takes shape because of our activity and therefore we and the object take shape together (Varela, 1999, 66). From these premises an important conclusion comes: if the basis of knowledge as representative process is in the perceptual-motor mechanism of interaction, it is not so important the fidelity of the representation/simulation of a virtual environment. The exchange of information in a virtual environment can be totally considered an exchange of information organism-environment.
In the evaluation of the cyber archaeological applications it is fundamental to know the epistemological commensurability. The increased information, its “density” constitutes the focus of a research path of virtual archaeology. In this assumption cyber archaeology is a system of communication and validation of the research approaches bottom-up and top-down; it is a *syllogis* of data and dynamic evidence not deductable by logics of feedback, that is behaviors able to generate other behaviors, actions making contexts and information. In the bottom-up approach we identity the operations of information input, in the top-down phases the actions of representation and mental patterning of information (Forte *et alii*, 2006).
Cyber Archaeology can represent today a research path of simulation and communication, whose ecological-cybernetic relations organism-environment and informative-communicative feedback constitute the core, but they have to be still fully investigated (Forte, 2007).
Yet if the cybernetic model is the focus, as example of interactive behavior, it is possible to study the relations between observers and models (both inter-connected). This aspect was in part discussed in the volume “Virtual Reality in Archaeology” (Forte, 2000): “It is useful, in fact, to notice that the major part of VA applications so far developed do not have important archaeological “contents”, nor, as would be worthwhile, respond to precise questions. Instead they tend to float in a generically popular and multimedial sphere, or they are used as technological exercises or as a means of rendering the archaeology more spectacular; completely separate from the research context and from the exegesis of the data. Noticeable gaps are represented by the fact that the models are not “transparent” in respect to the initial information (what were the initial data?) and by the use of the peremptory single reconstruction without offering alternatives (it could have been like this but we can also offer other models...).

Figure 3: The virtual reconstruction in transparency of the frescos of the villa overlapped to the laser scanner model.

The all too often missed occasions by VA are due to the clear and unbridgeable separation between the capacities of the archaeologists and computer scientists and to the inexistence of an interdisciplinary figure”. (Forte, 2000).
Two issues over all: transparency of data and peremptory of the reconstruction are not enough considered in the process of multidisciplinary research (Forte, Pescarin, Pietroni, 2006). The transparency of data is a crucial issue because it involves the validation process and all the data entry until to the final architectural modeling. Understanding all the work allows to increase the cybernetic “difference” in the learning activity between mind and ecosystem. To see in transparency means to verify the reliability of the work of virtual archaeology and to understand its development from the starting point to the final model. The issue then of the rigidity of the reconstruction (is there one reconstruction or many possible reconstructions?) can be solved in the relation of interaction between observer and model, namely in the dynamics of learning within the virtual ecosystem (Forte, 2007). The virtual reconstruction as research and communication process is always a selection between many possible reconstructions and it cannot represent ever the definitive solution for the archaeologist’s job.
Cybernetic archaeology should become mainly the workshop of scientific research, an active and measurable space where to compare datasets, models, hypotheses, archives, a cyber space of interactive knowledge.

2.1 Cybernetic models and reconstructions

A simple correspondence (equivalence) virtual archaeology=reconstruction of the ancient world seems, in some terms, reductive, or, otherwise, oversized, utopian. Reductive because it seems finalized to the methods of structural architectural recomposition and not to the study of processes and relations between architecture-environment-organisms. Utopian because reconstructing the ancient world is interesting as method, but not realizable in a single process. Finally the transmission of an interactive and cybernetic model should allow also the future communities of scientists to continue our work, correcting our errors and suggesting new archaeological interpretations.

In epistemological sense the ancient world cannot be reproduced and reconstructed, but in the attempt to recompose the context it is possible to codify the relations/affordances which the space-time has canceled. In short we could say that cyber archaeology is aimed to the construction of a spatial-temporal relations able to reconnect the territory with the "map", the archaeological landscape with the ancient landscape, following a validated and transparent methodological path.

The communication of any artifact or ecosystem depends on the transmitted and connected code. The reconnection of the relationship map-territory gives us the capacity of interpreting the past getting a major amount of information through the mutual interaction between observer and environment, where the same observer is part of the virtual ecosystem (Schroeder, 1997).

3. 3D INFORMATION

3.0 3D Environment

If our deeper knowledge of the environment is based from the perception of spatial coordinates and of the third dimension, a 3D digital ecosystem should be able to communicate a major amount of information and, mainly, to increase the dynamics of learning. The modality of perception and mental representation of the models contribute to the mediated knowledge of the world.

The Villa of Livia was fully documented by a time of flight's laser scanner: it means that a laser spot of a few millimeters makes almost in real time a model of the Villa. At the beginning the model is around several millions of points, then, after the optimization and decimation in meshes and polygons is a few thousands of points. From this analysis it is possible to understand that new methodologies of archaeological research return us an amount of data much greater than in the past. This involves a different ontological phase, diverse perceptual levels and complex forms of communication. What can we do with all this digital? What happens between representation and knowledge? How much are we influenced by aesthetic-perceptive properties of the model? Communication and information of a model depend substantially on the interaction, namely we have to imagine a dynamic process modified by the movement, light, perspective, geometry and from all the relations with the environment. For example in fig.4 we can see two different versions of a wall (with plaster and painted decoration) of the Villa of Livia. The high resolution model corresponds to a model of 46.145 polygons generated from a point cloud taken by a laser scanner, while model at low resolution is reduced to 1237 polygons with a normal mapping processing. The visual perception of the two models is very similar, but do they communicate the same kind of information? It depends on the final aim and representation: if we want to make a detailed analysis of the geometry of the model (structural calculations, measurements, volumes, etc.), the version with 1237 polygons would be not enough. On the contrary, if we explore the model in real time, this perception could be enough for a first interpretation.

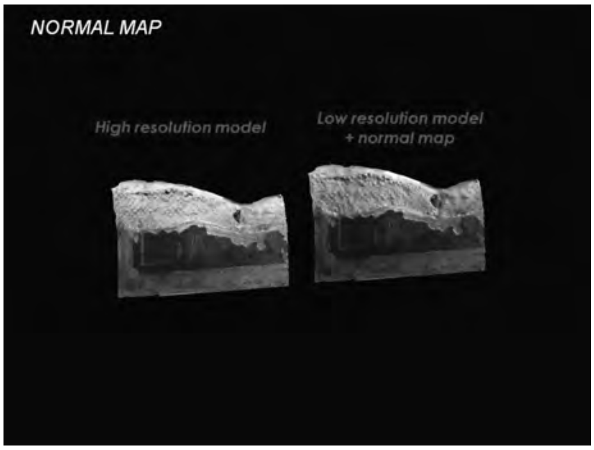


Figure 4: Normal map's application. Virtual reconstruction of a wall in high (left) and low (right) resolution

In conclusion, the perception in 3D spaces is a dynamic phenomenon and concerns firstly behaviors and effects. We list the main items:

- *Feedback.* Each action in the virtual space involves a result and a rule of learning.
- *Behaviors.* In the cyberspace it is possible to define pre-ordered and not pre-ordered events (for example the 3D navigation). Both categories enrich the virtual ecosystem, embodiment and capacities of learning.
- *Embodiment*. Ability to see the body as a place of knowledge processing in the dynamics of the virtual. The places of embodiment are also those of the hyper-real, of the augmented space, of the digital ecosystem.
- *Difference.* We learn through the difference: a difference generating a difference is an idea; a bit, that is an information unit (Bateson, 1979). The more is the difference between actor and ecosystem, the more is the capacity of exchange and communicates information. The representation in 3D creates a major difference in cybernetic sense; it means that interacting with datasets in 3D we develop a major exchange with the cybernetic ecosystem.
- *Space.* The 3D space is inter-connected and homogenizes relations and objects in the same scale and size.
- *Multisensoriality.* Virtual reality is multimodal and partially multisensorial (it is mainly based on audio-video). In any case even a partial involvement of our senses increases the perception of the three dimensions and characterizes the sense of place.
- *Light.* The 3D navigation develops the sense of embodiment, the sense of space and the environmental properties. Different light conditions need a more complex reading of information and augment the capacity of environmental learning.
- *Transparency.* The reconstructive process can be validated from a sequence of 3D worlds overlapping and spatially compatible.

- *Connectivity.* The spatial information in a three dimension multiplies its communication model in a conceptual network of links.
- *Accuracy.* The characterization of space depends on the spatial accuracy and on the abilities of representation and consumption of the models.
- *Cyber-realism.* Setting and sense of place are correlated with the qualities of photo-realism or from the expectations of the observer in the virtual environment. The expectations of realism increase the level of familiarization and embodiment in the virtual environment.
- *MUDs and social communication.* The agents within the system, for example avatars or subjective interactions, can learn through an unconscious imitation, following others' movements and by spatial sharing.

3.1 Cybernetic model

The cybernetic model of the Villa of Livia is a system of relations created by the real time interaction and navigation. It means that at theoretic level the cybernetic model does not have a preordered quantity of information, but it is progressively enriched by the explorations, integrating what is observing and what is observed (Forte, 2007).
The importance of the cybernetic model in comparison with the computational one is absolute, like the difference between logic and mathematics: the push to think of the cybernetics represents a real cultural model. For the cybernetics the information is the capacity of the organization level and complexity of a structure, in the sense that if a whole is random, it is not necessary to give some instruction for reproducing it (Wiener, 2001).
If the feedback constitutes the focal point of the informative dynamics, the description of the context is given from the relations. In a complex system the relations between elements are more important than the elements themselves (Forte, 2007).
The logic of a virtual reality system is similar to an anthill: each action can exchange a small amount of information with the system, but in holistic sense the sum of several actions makes a more intelligent and evolved exchange. The logic of the anthill can explain the holistic interpretation of the villa of Livia in the cyberspace. In the exploration of the space of the villa, room by room, area by area, we progressively arrive to recompose the logic and connective unit of the monument as a coherent and working structure.
One more important issue regards the criteria for selecting non verbal communication and not explicit codes. Many relations in the cyberspace do not have a name or a label, but transmit information in the dynamic of the system. The communication happens between movement, interaction and representation (Wiener, 1948). In the archaeological landscape, partially visible and readable in the modern landscape, the code is represented by an interpretation of the past aimed to the reconstruction of the ancient landscape (Forte, 2005).

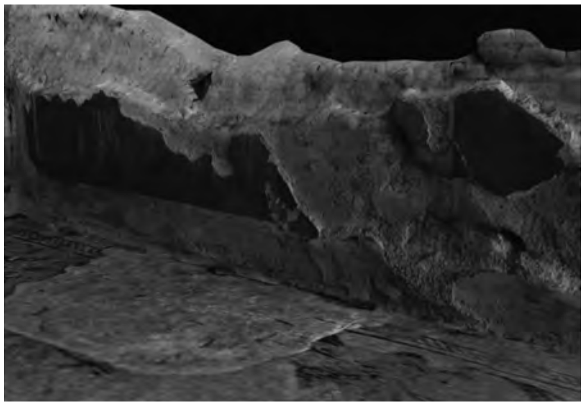


Figure 5: 3D visualization of a detail of a wall of the Villa, room 23, with plaster, frescos and chromatic components

This code is the map of the territory (Bateson, 1979), the interpretation key. Following these premises, the virtual reconstruction of the Villa of Livia is a simulation process: hence “the Villa” does not exist by itself, but the complex of potential and real relations linked with it exists, as *affordance*. The transparency of graphic materials used in the Virtual has allowed seeing trough the models, so that to have an easy structural comparison with the archaeological remains on site (figs.3, 7-8). In this way the virtual anastylosis is interpreted as integration of the remaining architectural models.
The capacity of learning according to dynamic models and the interaction between elements and environment is defined *enaction* (Mellet-d'Huart D., 2006). The enactive vision introduces the definition of an embodied mind in the environment; for this reason it is an appropriate approach to a virtual ecosystem. In fact there is a strong link between world and observer: "A history of structural coupling that brings forth a world. This is the term for the reciprocal process by which an observer educes unities from her medium within the limits of her phenomenology (i.e., as constrained by her embodiment) and the ontogenic coupling results in incremental regularization in the structure of the observer (her embodiment)” (Varela *et al*., 1991, 206). Then: "The fundament of an enactive account is not an objective ontological substrate, but the phenomenology of the individual defines enaction in terms of two intertwined and reciprocal factors: (1) the influence of an actor's embodiment in determining the trajectory of behaviors; and (2) the historical transformations which generate emergent regularities in the actor's embodiment”. These two aspects can be mapped onto two different usages of the English verb 'enact'. First is 'to enact' in the sense of 'to portray, to bring forth something already given and determinant of the present', as in a stage actor enacting a role. (Varela *et al*. 1991). The reciprocity of informative processes is the principle through which the observer is part of the virtual system, increasing its self-organizing capacity.
Each action in a virtual environment involves a feedback; the effect of this feedback is the perceptual-motor learning (bottom-up). In the case, for example, of a linear transmission of information (for example through a book), we have a symbolic-reconstructive learning (top-down). In cybernetic sense this mechanism can be described as in-out, or from the internal to the external environment, from the interaction to the learning.
In effect, the brain-training of the observer-actor, determined by the feedback of the system, allows an evolution to the use of the system with active and passive imitative processes. Active, when the observer learns from what he/she is doing, passive, when he/she learns from action of other users/observers.

4. CYBER RECONSTRUCTION

4.0 The Virtual Reconstruction of Villa of Livia

In the case of the Villa of Livia it is possible to access to the information's digital archive constituted by reconstructions, comparative models and graphic libraries. In short, the model is an open space aimed to grow and to be updated in the future, on the basis of further investigations on site or in post-processing. For what it concerns the issue of the reliability and congruity of the reconstruction three gradients (visualized with different nuances) have been conceived. The darkest nuance indicates a reconstruction which is totally scientific and reliable while the lightest indicates an evocative reconstruction that is based exclusively on generic cultural models of reference. In this way the virtual system is defined as a simulation environment and not as a simple virtual *maquette,* reproduction in scale of a hypothetic "original", just because this original cannot exists.
In fact, the creation of *maquettes* is closer to the idea of *replicas* than to the model of interactive simulation. The scientific coherence of the model in fact depends also from the faculty to distinguish the different ontologies of data: *in situ*, reconstructed, simulated, comparatives, dynamics, etc. It would be in fact too authoritative saying "this was the Villa of Livia in the I cent. A.D.", while the simulation enables the coexistence of different hypothesis and models of reconstruction especially in relation to the special context and to the landscape.
In practice the dynamics of simulation in a cybernetic process permits the combination of a high number of factors, behaviors, artifacts, ecosystems whose focus lies in the process and not in the single element or in the formalization of unique elaboration.
The research prospective of cyber archaeology is therefore of a holistic and constructivist type: the reality of information is in the perception, in the capacity to identify the possible realities not THE REALITY. The Villa of Livia, as a model of knowledge is segmented in different domains: the villa *in situ*, the villa through the sources and the excavation documentation, the villa and the landscape, the villa's reconstruction, the perception, the communication, the relations, the environment, all these and much more is the Villa of Livia Drusilla. The Villas of Livia therefore constitutes the ontology of information to interpret and communicate in reciprocity of intents of communication. A fundamental, I believe, mistake of virtual archaeology or maybe its original sin, was to separate the domains of knowledge and observation (what we know and we see today) from those of the hypothetic reconstruction, with the result of leaving visible and usable only the final state of the dialectic of interpretation.
For example the location of a site in the landscape, either in its original geo-context or in the relations with the ecosystem, multiplies the faculties of contextualizing the connection with other elements of the environment (figs.8-9), natural or artificial, as the parts of a monument are broadly speaking interconnected with its structure.
The methodologies of reconstruction in virtual archaeology, in particular with reference to the Villa of Livia can be classified schematically in this order:
Virtual Anastylosis: it deals with reconstruction of the ancient on an architectural and formal base in which the monumental space is privileged in respect to other possible simulations. In this case volumes and architectural forms are privileged in respect to materials, colours and textures. The Virtual Anastylosis can be also the first step to proceed to more complex reconstructions.
Evocative Models. In the evocative models the objective is to reconstruct by macro classifications, by comparative analyses without much attention to the relations with the data from fieldwork and to the spatiality of the information (fig.7). In this category are included the graphic 3D libraries, the serial contextualized architecture of landscapes and every generic modal but identifiable in the cultural attribution.

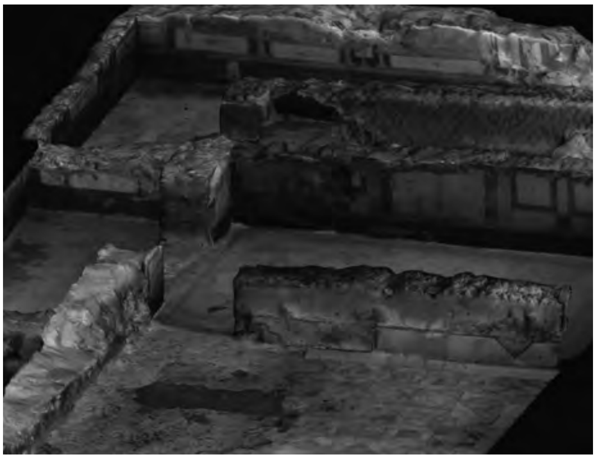


Figure 6. 3D model of the Villa of Livia by laser scanning data

Figure 7. Villa of Livia, 3D reconstruction
of the garden in the Republican age.

Hybrid models. They are models in which the reconstructed part (how the monument was in ancient times) integrates in transparency also the structures still preserved *in situ*. The hybridization is obtained from the coexistence of two architectural classes, real and reconstructed. In the model of the villa this hybridization makes easy the interpretation of the monumental structures (foundations and walls, fig.3).
Holistic reconstructed models. In this case the reconstruction integrates the architectural models, the textures and the furniture (fig.10). The simulation plans an integral reconstruction of the ancient villa.
Behaviors and organisms. They constitute the principal activities of avatars and agents: they can be active behaviors determined from users and passive behaviors identified as hypermedia links. For example an avatar-user meets a character in the virtual world which starts a movie or a tale.
Landscapes (figs.8-9). The artificial structures are fully integrated in the landscape and in the environment, whose physiography, vegetal coverage and ecological relations are

reconstructed. Natural and artificial landscapes are not separated domains, but they are part of the same ecosystem.
Finally, the cybernetic reconstruction of the Villa of Livia is characterized by the following features: transparency and hybridization of the models, affordances, reliability and validation of the reconstruction, geo-spatiality, behaviors, 3D, embodiment, MUD.

Figures 8-9: Villa of Livia: the archaeological landscape (top) today and the ancient Roman landscape (bottom). 1. Villa of Livia 2. Ancient Roman Flaminia road 3. Modern Flaminia road

5.0 Conclusions

The core of a cybernetic model in archaeology is the simulation relational process and the epistemological approach adopted. In this paper we have tried to redefine the role and the definition of the virtual archaeology as a cybernetic simulation process. The focus of this process would be not in the reconstruction itself, but in the multiple relations and "differences" produced by the interaction between users, environment and behaviors.
It is quite urgent therefore to plan that, in the mid of the digital era and with so many powerful tool of information processing, the scientific process in archaeology has to be review, mainly in the relationship between knowledge and communication. The importance of the new tools and technologies used in archaeology creates still unexplored ontologies: remote sensing data, laser scanning models, photogrammetric models, virtual models, simulation environments. All this produces an enormous amount of data, whose scientific content is difficult to understand. What are the relations between acquired and represented data? Which capacities of analysis, interaction and simulation? How much information does a cyber model communicate?
The case study of the Villa of Livia has created a remarkable amount of models related with the architecture, landscape, and ecosystem. It has integrated the detailed reconstruction of the archaeological landscape (the site today) with the simulation of the ancient landscape (the site in Roman times). This study has suggested new paths in the integration of field technologies, new models of study and communication, until to the virtual museum, the last step of this holistic interpretation. All this is aimed to define a diverse model of knowledge and communication, nomadic, open, accessible and finally definable as ecological digital process. The spatial sharing in a MUD space stimulates imitative and mutual information processes, catalyzing the cultural transmission.
It seems hence quite evident as the methodology of the archaeological research has to provide adequate epistemological tools for understanding the cognitive geometry of a cybernetic model. In the dynamics of interactive communication, all this complex of information is cyber archaeology and it belongs to an innovative process of reticular learning, where the observer is part of the ecosystem. We think one has to go towards a diverse formalism of scientific research in archaeology, rethinking the information domain.
In the reticular learning which is distributed through dynamic and interactive models, the cybernetic frame moves from the flat area of the display to embrace the environment and the observer in a diverse cognitive and perceptive logic, maybe still to be defined; but it is there, close to the margins of chaos, that the knowledge starts.

Figure 10: Villa of Livia, Southern part, room 6.

Acknowledgments

The Virtual Museum of the Ancient Via Flaminia was supported by Arcus spa and managed by CNR-ITABC (scientific direction) and National Roman Museum in Rome. All this activity was the result of my former job as director of the Virtual Heritage Lab of CNR-ITABC. A special thank is due to all the extraordinary personnel of the Virtual Heritage Lab with which I spent my last two years of research in Italy.

References

Antinucci, A., 2004, Comunicare il museo, Laterza, Roma, 2004.

Baudrillard J.. 1994, Simulacra and Simulation, Ann Arbor: University of Michigan Press, 1994.

Bateson, 1967, "Cybernetic explanation", in SEM, 410.

Bateson, 1972, *Steps to an Ecology of Mind* , San Francisco, Chandler Press.

Bateson G., 1979, *Mind and Nature. A Necessary Unit*, Dutton, New York.

Biocca F. 1997, The cyborg's dilemma: Progressive embodiment in virtual environments, Journal of Computer-Mediated Communication, vol. 3, n. 2, 1997.

Forte, M. 1997, (ed. by) *Virtual Archaeology*, (forward by Colin Renfrew) Thames & Hudson Ltd, 1997.

Forte, M., 2000, *About virtual archaeology: disorders, cognitive interactions and virtuality*, in Barcelo J., Forte M., Sanders D., 2000 (eds.), Virtual reality in archaeology, Oxford, ArcheoPress (BAR International Series S 843), 247-263.

Forte M., 2003, *Mindscape: ecological thinking, cyber-anthropology, and virtual archaeological landscapes*, in "The reconstruction of Archaeological Landscapes through Digital Technologies" (eds. M.Forte, P.R.Williams), Proceedings of the 1st Italy-United States Workshop, Boston, Massachussets, USA, November 1-3, 2001, BAR International Series 1151, Oxford, 2002, 95-108.

Forte M., 2005, *A Digital "Cyber" Protocol for the Reconstruction of the Archaeological Landscape: Virtual Reality and Mindscapes* in Recording, Modeling and Visualization of Cultural Heritage (eds: E.Baltsavias, A.Gruen, L.Van Gool, M.Pateraki) Published by Taylor & Francis / Balkema ISBN 0 415 39208 X, 339-351, 2005.

Forte et alii, 2006; M.Forte, S.Pescarin, E.Pietroni, C.Rufa, 2006, *Multiuser interaction in an archaeological landscape: the Flaminia Project*, in (M.Forte, S.Campana, eds.by) *From Space to Place,* Proceedings of the 2nd International Conference on Remote Sensing in Archaeology, Rome, December 4-7, 2006, BAR International Series 1568, Archaeopress, Oxford, 2006, 189-196.

Forte, M, Pescarin, S. Pietroni, E., 2006, *Transparency, interaction, communication and open source in Virtual Archaeology*, in (M.Forte, S.Campana, eds.by) *From Space to Place,* Proceedings of the 2nd International Conference on Remote Sensing in Archaeology, Rome, December 4-7, 2006, BAR International Series 1568, Archaeopress, Oxford, 2006 535-540.

Forte, M., 2007, *Ecological Cybernetics, Virtual Reality and Virtual Heritage*, in "Theorizing Digital Cultural Heritage. A Critical Discourse" (Edited by Fiona Cameron and Sarah Kenderdine), MIT Press, Cambridge, MA, 389-407.

Forte M., 2008 (ed.), La Villa di Livia. Un percorso di ricerca di archeologia virtuale, L'Herma, Rome, 2008.

Gallese, V. 2005, Embodied simulation: From Neurons to Phenomenal Experience, "Phenomenology and the cognitive sciences", 4, 23-48.

Gibson, J. J., 1999. *Un approccio ecologico alla percezione visiva* (Il Mulino: Bologna).

Ingold, T., 2000, *The perception of the Enviroment. Essays in livelihood, dwelling and skill*, London and New York, Routledge.

Korzybski A., 1941, *Science and Sanity*, Science Press, New York, 1941.

Maturana, H, Varela, F., 1980, Autopoiesis and Cognition: the Realization of the Living, Boston Studies in the philosophy of science, Cohen, Robert S., And Marx W. Wartofsky (eds.), vol. 42, Dordecht (Holland): D. Reidel Publishing Co., 1980.

Maturana, H, Varela, F., 1992, The Tree of Knowledge: the Biological Roots of Human Understanding, Boston: Shambhala, 1987, (Revised Edition: same publisher, 1992).

Mellet-d'Huart D., 2006, A Model of (En)Action to approach Embodiment: A Cornerstone for the Design of Virtual Environments for Learning, in Win W. & Hedley N., Eds. Journal of Virtual reality, special issue on education. Springer London. Volume 10, Numbers 3-4 / December, 2006. Pp. 253-269. 2006.

Morganti, F. Riva, G. 2006, Conoscenza, comunicazione e tecnologia: Aspetti Cognitivi della RV, Milano: LED

Piaget, J. (1980). *Adaptation and Intelligence*. London: University of Chicago Press.

Richardson, A. E., Montello, D. & Hegarty, M. 1999, Spatial knowledge acquisition from maps, and from navigation in real and virtual environments*, in Memory & Cognition*, *27*, 741-750.

Schroeder, R., 1997, *Networked Worlds: Social Aspects of Multi-User Virtual Reality Technology*, Sociological Research Online, vol. 2, no. 4.

Taylor, M.C. 2005, Il momento della complessità. L'emergere di una cultura in rete. Codice edizioni: Torino.

Varela *et al*., 1991, Varela, F., Thompson, E. - Rosch, E. *The Embodied Mind. Cognitive Science and Human Experience*, MIT Press, Cambridge, 1991.

Varela F.J., 1999, "Quattro linee guida per il futuro della conoscenza", in *Argonauti nella Noosfera. Mente e cuore verso nuovi spazi di comunicazione,* Vol. 2, strutture ambientali n.118/dicembre 1999, Atti della XXV ed. delle Giornate internazionali di studio promosse dal Centro Ricerche Pio Manzù.

Watzlawick, P. (ed.) (1985) The invented reality. New York, Norton.

Wiener, N., 1948, *Cybernetics, or control and communication in the animal and the machine.* Cambridge, Massachusetts: The Technology Press; New York: John Wiley & Sons, Inc., 1948.

Data Modeling and Infrastructure for Cultural Heritage Applications

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

A DATA BASE SYSTEM FOR MANAGING INFORMATION CONCERNING HISTORICAL MORTARS

I. Papayianni , V. Pachta *, K. Iliadou

Laboratory of Building Materials, Dept. of Civil Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece – (papayian, vpachta)@civil.auth.gr

KEY WORDS: Binding agents, Data base, Historic mortars, Holistic analysis, Renderings

ABSTRACT:

Mortars have played a very significant role in the technological evolution of construction, from the prehistoric era until the 20th century. The types of raw materials used, the proportioning of the mixtures, the techniques of application and their resistance to weathering are some of the topics concerning the Building Technology. A large number of mortar samples, more than 1500, from all historic periods of the Hellenic land, were systematically analyzed in the Laboratory of Building Materials AUTH, the last 15 years, according to the holistic approach which was developed there. A great number of results should be considered and evaluated from different points of view, so as the philosophies of old technology and the knowledge of the historical, archaeological, architectural or even ethnological interest to be revealed. In order to evaluate easily the properties of mortars of each era, a flexible data system was designed, which could include all results from the samples' analysis. The data input was designed to be easily comprehended by any user, while the presentation of all results is based on the interest of the user. It can refer to the mortars of a specific monument or give general information regarding the types and properties of mortars used during a historic period. The basic aim of the data system was to concentrate all archives of the Laboratory in a unique file, however through its use it seems to be a great tool for any scientist engaged to the field of restoration, since in a very easy and flexible way one can be capable of having and handling a great source of information regarding historic mortars' analysis.

1. INTRODUCTION

1.1 General information

Mortars (named also in literature as emplecton, titanos, cocciopesto opus cementitium, stucco, courasani, lykium) is a diachronious building material which dated from the first effort of man to built a dwelling (Papayianni 1994).

The primitive man had even from the birth of humanity tried to create a dwelling in order to protect himself and his belongings. The types of structures usually had to fulfill the specific requirements of each era, while the types of the materials used were defined by the availability of raw materials as well as the role and form of the constructed element. The building materials as well as the constructive techniques changed through time, conserving however the basic principles of each society: the stability of the structure to provide protection from the environmental factors and the functionality of the enclosed spaces for the needs of the residents. (Papayianni, Pachta, 2005)

As binding materials, mortars were firstly used to joint ancient earthen blocks and their stones and bricks. During the Antiquity mortars were also found to be used as surface's plasters (of walls, floors or roofs) of houses, reservoirs or marine structures.

The technology of building, developed in parallel with society and culture. Some excellent samples of construction constitute up to nowadays the cornstones of human civilization.

1.2 Types of historic mortars

As far as the bonding materials (mortars) is concerned, the main factor of their production and use was, apart from the chronical and local characteristics of each era and area, their role in the structure. The main categorization of mortars, due to their use was (Papayianni 1994):

1. Joints of masonry (for bonding pieces of solid pieces of stone or brick)
2. Renderings (internal, external, of roofs)
3. Floorings
4. Roofing
5. Substratum of mosaics / frescos
6. Imitations of architectural members

1.3 Historic evolution of binding agents

The types of mortars used through time, were based on the technological awareness of each era, as well as the local availability of raw materials. The main constituents of historic mortars from prehistory until nowadays can be classified in:

1.3.1 Clay: It refers to the first material used in construction, due to its abundance and to its facile workability. The main applications of clay are firstly located during the neolithic era and were related to the manufacture of mud-mortars and mud-bricks. Apart from the technological evolution in construction, mud-mortars continued to be a basic bonding material, especially for rough constructions until the 20th century (fig. 1). It should be stated that even today a great percentage of earth's population still lives in adobe houses.

* Corresponding author.

Figure 1: Typical earth –block masonry of the 19th cent. in Veroia, Greece (Papayianni, 2006).

1.3.2 Lime: Although Greeks knew about lime even from the pre-historic times they used it as a bonding agent not until the 3rd cent. B.C. The main use of lime during antiquity, was to produce plasters with which walls, floors and roofs were covered, in order to be protected from humidity and rainwater. From antiquity until nowadays, lime is considered to be the most popular building material.

1.3.3 Gypsum: Gypsum was already known from the era of the Egyptian civilization. Its application in masonries, usually referred to its addition in lime-based mortars and especially renderings, in order to increase specific properties, such as heat-insulation, fire-insurance, workability and plasticity. (Papayianni, Pachta, 2005).

1.3.4 Pozzolana: Although pozzolana was considered to be a roman achievement in constructions, its use was identified in great civilization even from the prehistoric times (Minoan Crete, Hirokoitia Cyprus e.t.c.). Until the end of the 19th cent. all large projects (ports, isthmus of Korinthos, the Suez Canal) were produced by mortars and concretes which contained pozzolan and specifically santorine earth as a basic constituent. (Papayianni, Pachta, 2007). (Fig.2)

Figure 2: The archaeological site of Olynthos (5th cent. B.C.) the mortars taken from there consist of a high percentage of pozzolanic material.

1.3.5 Brick dust: It was firstly used in Hellenistic and Roman mortars (cocciopesto), mixed with lime, in order to increase the hydraulic properties of the mixture. During the Byzantine era brick dust was a main binding agent which reinforced lime mortars (Fig.3), colored the mixtures and preserved them from the negative results of humidity.

Figure 3: Colored lime-mortar containing brick dust from Hagia Sophia Thessaloniki (dated from the Byzantine period, 7th cent. AD).

2. AIM AND USABILITY OF THE DATA SYSTEM

2.1 Aim

The Laboratory of Building Materials, AUTH, has already a large experience in the analysis of historic mortars and the proposal of suitable repair materials for intervention purposes. The analysis is based on the holistic approach, developed in the Laboratory, which consists of a series of tests aiming in examining the physico-mechanical and chemical composition of each sample and strategic design of mortars to fulfil compatibility requirements.

More than 2000 samples of historic mortars from all historic eras and various monuments of the Hellenic land, were analyzed during the last 15 years in the Laboratory. As a result the archives have been enriched with a treasure of information regarding the types of mortars used in each era, as well as their technological characteristics.

In order for all this information to be easily approached and managed, a data system which would contain all the results form the analysis, turned to be an intense need. This data system should be designed in a rather flexible way, in order for all information to be easily handled by the user and lead to general conclusions based on the comparison of the results.

2.2 Usability

The basic aim of the data system was to concentrate all archives of the Laboratory in a unique file, in order for all researchers working in projects of restoration in the Laboratory, to be capable of collecting information from the archives. However, through its use, it seems that this data base can become a great tool for scientists of different discipline, engaged to the field of restoration.

Through the rich archives of the Laboratory one can make a quick retrospection in the types of materials used in each era, or even see results of a specific monument in which maybe he is interested in. This system could be also connected to other archaeological-architectural data bases or even sites, in order to give more detailed information regarding history, ethnology, archaeology, architecture, social and economic life of the area.

3. DESIGN AND APPLICATION OF THE DATA SYSTEM

The data system was realized by using the Microsoft Office Access, 2003, which is a program capable of having and managing a great volume of information. The specific design of the base, according to the needs of the user is the basic step, since in a well-designed base one can be sure of a flexible and efficient usage. In parallel, this system has the capability to be extended, in order for more information to be added in the future.

The data base consists of two basic elements: tables and forms. In the tables one can register all data, while the form is the way of their presentation.

3.1 Tables

The data base consists of four types of tables:

1. The table of the general historic eras in which all samples belong to (Fig. 4). All historic periods were defined according to literature (Bouras, 1999).

1 Περίοδοι : Πίνακας

Αναγνωριστικό	Ιστορική Περίοδος
1	Προϊστορική
2	Αρχαιότητα
3	Βυζαντινή
4	Μεσαιωνική
5	Οθωμανική
6	Σύγχρονη
ματη Αρίθμηση)	

Figure 4: Table of the historic eras

2. The table of the specific historic periods (Fig.5).

2 Ιστορικές υποπερίοδοι : Πίνακας

Κωδ υποπεριοδου	Περίοδος	Υποπερίοδος
1	Προϊστορική	Μινωϊκή
2	Προϊστορική	Κυκλαδίτικη
3	Προϊστορική	Μυκηναϊκή
4	Αρχαιότητα	Αρχαϊκή
5	Αρχαιότητα	Κλασσική
6	Αρχαιότητα	Ελληνιστική
7	Αρχαιότητα	Ρωμαϊκή
8	Βυζαντινή	Παλαιοχριστιανική
9	Βυζαντινή	Πρώιμη Βυζαντινή
10	Βυζαντινή	Μέση Βυζαντινή
11	Βυζαντινή	Ύστερη Βυζαντινή
12	Μεσαιωνική	Μεσαιωνική
13	Οθωμανική	Οθωμανική
14	Σύγχρονη	19ος αιώνας
15	Σύγχρονη	20ος αιώνας
ιυτόματη Αρίθμηση)		

Εγγραφή: 1 από 15

Figure 5: Table of the specific historic periods

3. The tables of all types of mortars (structural, renderings, substratum of mosaics, frescos, roofing, all types) (Fig.6).

Ιστορική Περίοδος
Κονίαμα δόμησης
Επίχρισμα
Υπόστρωμα τοιχογραφίας
Υπόστρωμα ενταίχιου ψηφιδωτού
Υπόστρωμα επιδαπέδιου ψηφιδωτού
Υπόστρωμα δαπέδου
Οροφής
Όλα

Figure 6: Table of the types of mortars

4. The tables of all monuments where the samples of mortars come from (Fig.7).

Κωδ Μνημείου	Κωδ υποπεριόδου	Υποπερίοδος	Μνημείο	Περιοχή	Χρονολογία κατασκευής
62	7	Ρωμαϊκή	Αψίδα Γαλερίου	Θεσσαλονίκη	
31	7	Ρωμαϊκή	Γαλεριανό συγκρότημα	Θεσσαλονίκη	3ος αιώνας
64	7	Ρωμαϊκή	Ρωμαϊκή Αγορά	Θεσσαλονίκη	
27	7	Ρωμαϊκή	Συγκρότημα Βαγονιών	Νικόπολη	
42	7	Ρωμαϊκή	Τείχη Διοκλητιανούπολης	Άργος	
32	8	Παλαιοχριστιανική	Ι.Ν. Παναγίας Δροσιανής	Νάξος	6ος ή 7ος αιώνας
38	8	Παλαιοχριστιανική	Κάστρο Μυτιλήνης	Μυτιλήνη	483 - 565 μ.Χ.
67	8	Παλαιοχριστιανική	Κεντρική Βασιλική	Δίον	
60	10	Μέση Βυζαντινή	Ι.Μ. Αγίου Παύλου	Άγιο Όρος	9ος αιώνας
25	10	Μέση Βυζαντινή	Ι.Μ. Δαφνίου	Αττική	11ος αιώνας
57	10	Μέση Βυζαντινή	Ι.Μ. Δοχειαρίου	Άγιο Όρος	10ος αιώνας
46	10	Μέση Βυζαντινή	Ι.Μ. Ιβήρων	Άγιο Όρος	10ος αιώνας
45	10	Μέση Βυζαντινή	Ι.Μ. Μεγίστης Λαύρας	Άγιο Όρος	963 μ.Χ.
56	10	Μέση Βυζαντινή	Ι.Μ. Ξηροποτάμου	Άγιο Όρος	10ος αιώνας
43	10	Μέση Βυζαντινή	Πρωτεΐο	Άγιο Όρος	10ος αιώνας
55	11	Ύστερη Βυζαντινή	Ι.Μ. Γρηγορίου	Άγιο Όρος	14ος αιώνας
34	11	Ύστερη Βυζαντινή	Ι.Μ. Κοιμήσεως Θεοτόκου	[illegible]	
61	11	Ύστερη Βυζαντινή	Ι.Μ. Κουτλουμουσίου	Άγιο Όρος	13ος αιώνας
48	11	Ύστερη Βυζαντινή	Ι.Μ. Παντοκράτορος	Άγιο Όρος	14ος αιώνας
49	11	Ύστερη Βυζαντινή	Ι.Μ. Σιμωνόπετρας	Άγιο Όρος	13ος αιώνας
53	11	Ύστερη Βυζαντινή	Ι.Μ. Χιλανδαρίου	Άγιο Όρος	13ος αιώνας
37	11	Ύστερη Βυζαντινή	Ι.Ν. Αγίου Δημητρίου	[illegible]	
22	11	Ύστερη Βυζαντινή	Παναγία Κρήνα	Χίος	13ος αιώνας
13	12	Μεσαιωνική	Αγία Αικατερίνη	Ρόδος	14ος αιώνας
14	12	Μεσαιωνική	Αρχοντικό	Ρόδος	
15	12	Μεσαιωνική	Ιπποτικό	Ρόδος	
40	12	Μεσαιωνική	Κάστρο Αγίας Μαύρας	Λευκάδα	1300
16	12	Μεσαιωνική	Κατάλυμα Ισπανίας	Ρόδος	14ος αιώνας
18	12	Μεσαιωνική	Παναγία Μπούργκου	Ρόδος	11ος αιώνας
19	12	Μεσαιωνική	Τέμενος Σουλεϊμάν	Ρόδος	1522 μ.Χ.

Εγγραφή: 1 από 44

Figure 7: Table of all monuments

5. The table where all the results from the mortars' analysis are being registered. In order for the information to be more specific there are as many tables as the types of mortars confronted (structural, rendering, substratum of mosaics, of fresco, roofing) (Fig. 8)

Αναγνωριστικό	Κωδ απο	Μνημείο	Κωδ αρ δείγματος	CaO	SiO2+Al2O3+Fe2O3	Αδιάλυτα συστατικά (% κβ)	Πίνακας
[illegible]	7	Αψίδα Γαλερίου	7	27,9	35,79		
199	7	Αψίδα Γαλερίου	9	33,6	42,17		
66	7	Γαλεριανό συγκρότημα	2	41	17,37	9,6	
67	7	Γαλεριανό συγκρότημα	3	41,5	11,24	5,4	
68	7	Γαλεριανό συγκρότημα	4	43,4	13,27	11	
69	7	Γαλεριανό συγκρότημα	5	41,3	11,64	13,7	
70	7	Γαλεριανό συγκρότημα	7	42,5	18,28	14,3	
215	7	Ρωμαϊκή Αγορά	10	24,8	30,42		
196	7	Ρωμαϊκή Αγορά	8	31,4	30,7		
198	7	Ρωμαϊκή Αγορά	9	25	34,25		
119	7	Τείχη Διοκλητιανούπολης	10	42,2	22,87	55,8	
84	8	Ι.Ν. Παναγίας Δροσιανής	10	30,9	40,56	60,2	
85	8	Ι.Ν. Παναγίας Δροσιανής	11	38,8	23,49	39,9	
86	8	Ι.Ν. Παναγίας Δροσιανής	12	40,7	23,93	60	
78	8	Ι.Ν. Παναγίας Δροσιανής	1α	33	33,17	75,9	
79	8	Ι.Ν. Παναγίας Δροσιανής	1β	29	44,28	86,9	
80	8	Ι.Ν. Παναγίας Δροσιανής	3	37,5	23,25	40,7	
81	8	Ι.Ν. Παναγίας Δροσιανής	4	27,4	41,24	76,1	
82	8	Ι.Ν. Παναγίας Δροσιανής	7	36,3	24,63	23,8	
83	8	Ι.Ν. Παναγίας Δροσιανής	9	44,2	16,49	26,3	
109	8	Κάστρο Μυτιλήνης	1	68	48,54	22,6	
110	8	Κάστρο Μυτιλήνης	5	67	54,25	59,4	
111	8	Κάστρο Μυτιλήνης	M2	0,66	49,49	32,8	
183	8	Κεντρική Βασιλική	1	34,8	26,3		

Εγγραφή: 1 από 210

Figure 8: Table of all results per sample of mortar

3.2 Forms

In order for all the results to be easily presented, a series of forms were designed. These forms are connected with relevant tables and take all data from them. They can be classified in:

1. The front page, where one can choose inbetween the alternatives: presentation of all results, data processing, addition of results, data analysis. By choosing one of these alternatives the user goes to the specific pages. (Fig.9)

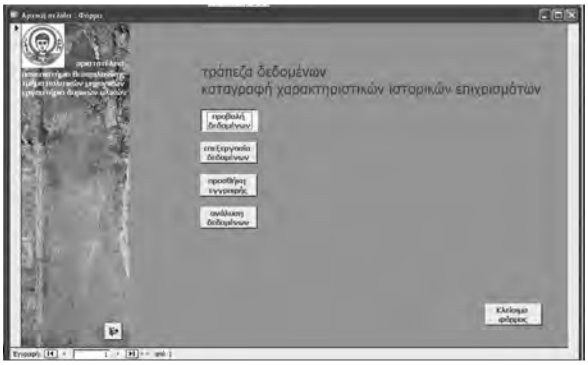


Figure 9: The front page of the data system

2. The form of the historic eras and specific periods which takes all data from the relevant table (Fig. 5). By choosing one specific period, one can go through all the mortar samples from monuments of this period (Fig. 10).

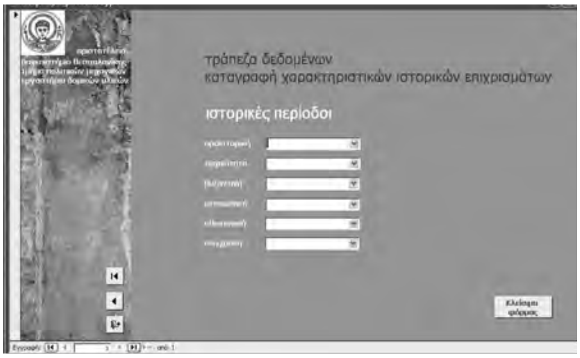


Figure 10: The form with all historic periods

3. The form of a specific historic period, from where one can choose the monument, as well as the type of mortars he is interested in (Fig.11, 12). In case there is not a specific type of mortar in a monument (p.e. rendering) then a form indicates it (Fig. 13).

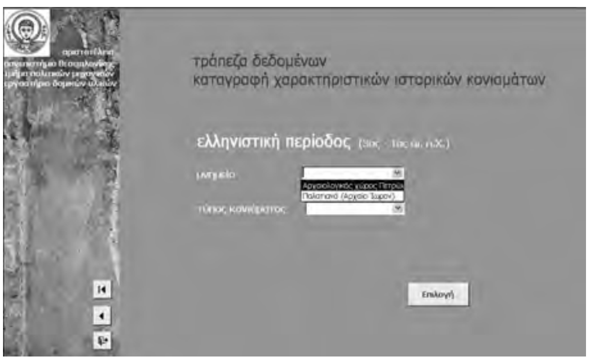


Figure 11: The form of a specific historic period (Hellenistic), with all monuments and types of mortars

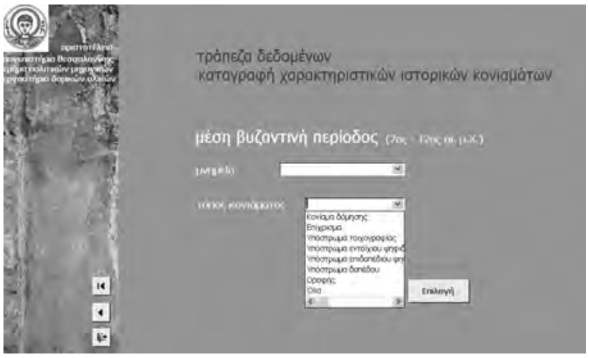


Figure 12: The form of a specific historic period (middle Byzantine), with all monuments and types of mortars

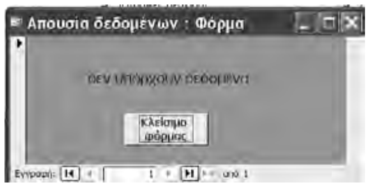


Figure 13: The form indicating that there is no data for a specific type of mortar in a monument

4. The form presenting all the results from the analysis of a specific sample of mortar from a monument (Fig.14). This form takes all data from the relevant table (Fig.8) and the information presented concerns:

- Information regarding the monument (name, location, period of construction, photos).
- General information regarding the sample (code number, type, sampling area, whether it is authentic or regards a subsequent constructional phase or contemporary intervention, photos or designs from the sampling area, macroscopic and microscopic photos).
- All results from the analysis which can be classified as:
- Stereoscopic analysis, regarding the structure of the sample, the presence of pores and cracking, the types of aggregates, enclosed materials, other observations.
- Physico-mechanical regarding its compressive strength and porosity, Apparent specific Gravity.
- Chemical, regarding its proportion in $Ca(OH)_2$ and argilo-silicious components, as well as all results from its chemical composition (percentage of oxides, soluble salts).
- Granulometric analysis (percentages of aggregates, granulometric curve).

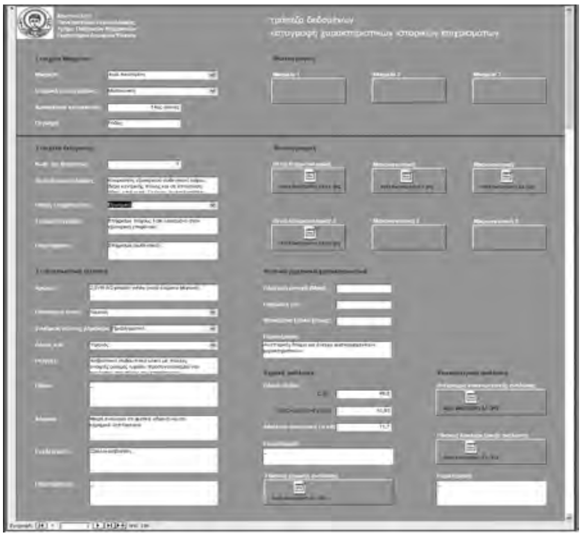


Figure 14: The form with all the results from the analysis of a specific sample.

4. MANAGEMENT AND ANALYSIS OF ALL DATA

The basic advantage of this data base is that the user can separate the information he is interested in (p.e. the results from a specific type of mortar of a monument) or take general information regarding a whole category of samples (p.e. results regarding the renderings of a specific historic period) and statistically analyze them.

The statistic analysis of the results is realized through the queries.

4.1 Queries

The queries enable the user to track down and handle the information he is interested in. These queries can be based in one or more relevant tables and gather data from them.

This alternative was multiple used in the data system, per example for the determination of the composition of renderings from all historic periods in lime (percentage of $Ca(OH)_2$) and argilosilicious components (SiO_2 + MgO + Fe_2O_3). In addition, queries were used for the determination of the average of the physico-mechanical properties of all types of mortars and of all the historic periods.

4.2 Statistic analysis of results

From the analysis of all data, it was concluded that the user is capable of handling a great source of information regarding all types of mortars from all historic periods, taken from monuments of Greece. The statistic analysis of all this data can be easily achieved, since all results gained from the queries are possible to form graphics for their better examination.

The major goal from this achievement was that through the analysis, many conclusions arise, regarding several topics relevant to the technological evolution of historic mortars, such as:

- Structural, chemical, physico-mechanical properties.
- Types and proportions of raw materials used (even in comparison with the specific characteristics of each era and the local physiognomy of each monument).
- Morphological characteristics.
- Techniques of application.
- Specific observation regarding each historic era.

4.3 Conclusions derived from the analysis

From the analysis of the data concerning renderings of all historic period it was concluded that the composition of all renderings of all historic periods is based in lime, since the proportion of $Ca(OH)_2$ is rather high (>40%). This conclusion certifies the prediction that all renderings are lime-based.

In Figure 15, the diagram compares the proportion of $Ca(OH)_2$ of renderings of all eras.

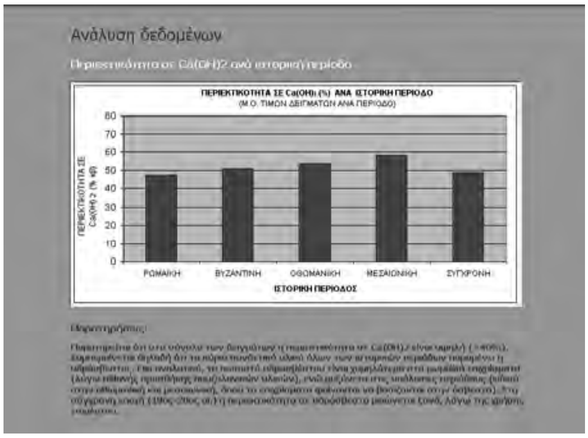


Figure 15: The proportion of $Ca(OH)_2$ of renderings of all eras

According to the diagram, the proportion of lime is lower in roman renderings (probably due to pozzolanic components), while it is increased in the rest periods (especially during ottoman and medieval, where renderings are lime-based). During the 19^{th} and 20^{th} century the proportion in lime is decreased (due to the use of cementitious materials).

In contradiction, according to Figure 16, the percentage of argilosilicious components (SiO_2+Al_2O_3+Fe_2O_3), is rather increased in roman renderings (where there is a systematic use of pozzolanic materials). In addition this percentage is also increased in Byzantine times (intensive use of brick dust) and in the 20^{th} century (use of cementitious materials).

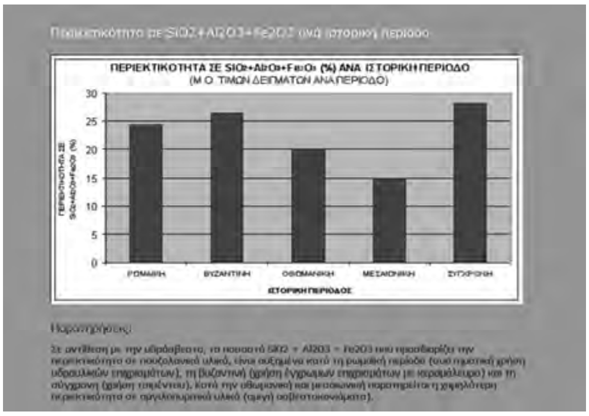


Figure 16: The proportion of argilosilicious components of renderings of all eras

From the comparison of the above diagrams it is concluded that in all historic periods prevail lime based renderings. According to Figure 17 the percentage of the ottoman lime based renderings is the higher one (96.7%) and follow the medieval ones (92.3%), the Byzantine (87.4%), the 19th-20th century (86.9%) and finally the roman renderings (72.7%).

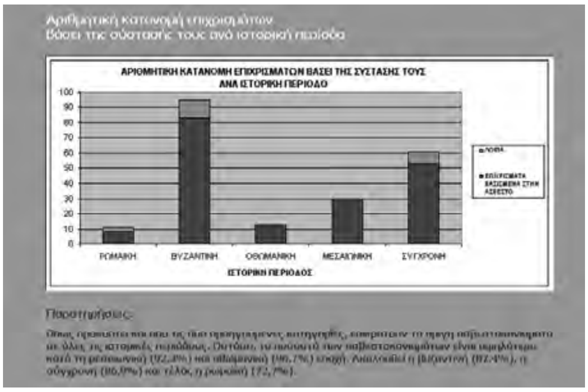


Figure 17: Lime based renderings per era

From the analysis of the properties of renderings there has been made a comparison of their physical properties (porosity, apparent specific gravity). In the total number of the samples the porosity is high (>20%), with the apparent specific gravity to be rather low (1.6-1.9). (Fig. 18)

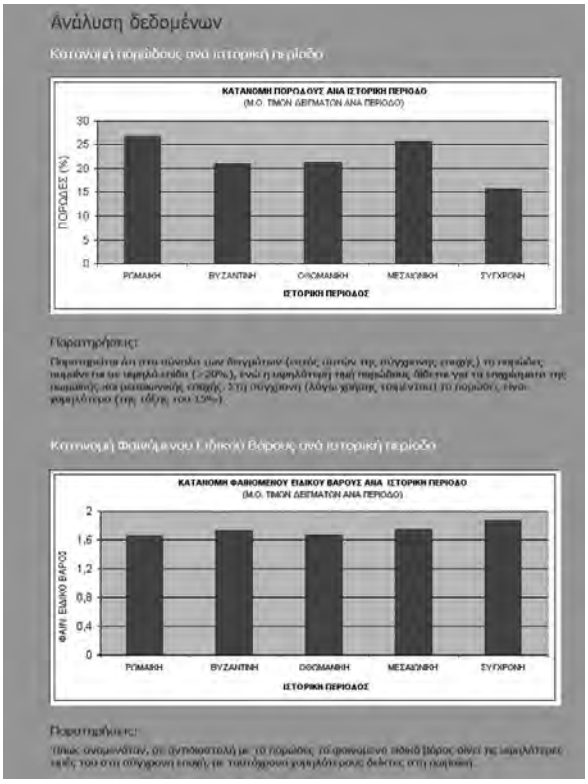


Figure 18: Physical properties of renderings of all historic eras

In accordance to all these graphics, several conclusions were made regarding all types of mortars. These conclusions made easy the comparison of all types of mortars of every historic period.

5. EXPLOITATION OF THE DATA SYSTEM

The rich archives of the Laboratory of Building Materials AUTH, regarding the analysis of historic mortars are considered to be rather valuable in national as well as in international level. The high experience in the methods of the samples' analysis, as well as the great number of the tested mortars, renders these archives capable for an objective study of building materials through time.

Until now, a long time was necessary for the study of the archives and the management of all information. However, the data system has made this attempt easy by anyone interested in this field, since all results are presented is a clear and flexible way. Even if someone is not familiar with the terminology of the scientists working in the Laboratory can gain important information by using the system.

Since the data system is still in progress, not all the ways of its exploitation are already clarified. However, its rich data in comparison with the easy way of its managing can lead to specific exploitation ways, such as:

1. Better organization of all archives of the Laboratory in a unique file.
2. Easy management of all data.
3. Pilotic use of the data system by not experienced users, in order to see disadvantages and difficulties confronted.

4. General use of the system by researchers engaged to the field of restoration, in order to see its impact and the help it can offer in technological queries.
5. Connection of the system to other national data bases concerning archaeological, architectural, technological themes.
6. Connection of the system with other international scientific bases, in order for more general conclusions regarding building materials of all regions and periods to be achieved.
7. Developing additional tools for facilitating users, such as:

- Terminology / Glossary
- Characteristics of historic periods
- Map of archaeological interest
- development of a site of communication or answering questions
- Establishment of an administrating center.

6. CONCLUSIONS

The basic aim of the data system was to concentrate all archives of the Laboratory in a unique file in order for their management to be easily achieved. However through its progress and use it seems that it can become a great tool for any scientist engaged in the field of restoration, since in a very easy and flexible way one can be capable of having and handling a great source of information regarding historic mortars' analysis.

Taking into account that the international information regarding historic mortars is rather limited, as well as the fact that there is not a regulation nowadays regarding repair mortars, it can be assumed that this data system can become a major help in the field of the technological evolution of historic mortars.

REFERENCES

References from Conferences:
Papayianni I., 1994. *Durability Lessons from the Study of Old Mortars and Concretes*, Lecture presented at ACI P.K. Mehta Symposium on Durability of Concrete, Nice France, pp.1-20.

Papayianni I., 2006. *Restoration of Earth Block Buildings in Vernacular Architecture. Materials and Techniques for their Repair*". ECTP – European Construction Technology Platform, Construction aspects for Heritage Protection – Research needs, October 15th – 18th 2006, Croatia.

Papayianni I., Pachta V., 2005. *The evolution of the binding agents' technology,* International Seminar, 'Theory and Practice of construction: knowledge, means, models', Edizioni Moderna, Ravenna 27-29 October, pp. 305-315.

Παπαγιάννη Ι., Πάχτα Β., 2007. *Ιστορία της χρήσης των Φυσικών Πουζολανικών υλικών στη δόμηση.* 1ο Εθνικό Διεπιστημονικό Συνέδριο «Ιστορία των Δομικών Κατασκευών», Ξάνθη, 29 Νοεμβρίου - 1 Δεκεμβρίου 2007.

References from Books:
Bouras Ch., 1999, *History of Architecture,* 1st, 2nd Volume, Melissa Editions.
Davey, N., 1961. *A history of Building Materials*, London: Phoenix house.

Lawrence A.W., *Greek Architecture,* Penguin Books, G. Britain, 1962. pp.3-5.

Orlandos A.K., *The Building Materials of the Ancient Greek,* Vol. I, Vol. II, Athens 1958. pp.Vol. I: 46-66, Vol. II: 65-85.

ACKNOWLEDGEMENTS

The authors express their appreciation to the Hellenic Cement Industry Association, for the support of the Project "Study of the technological evolution of mortars, concretes", which is realized in the Laboratory of Building Materials, AUTH.

SAINT CLASSIFICATION IN BYZANTINE ART

K. Raftopoulos [a], P. Tzouveli [a], K. Ntalianis [a], S. Kollias [a], D.Kalomoirakis [b], N. Fyssas [b], G. Foukaneli [b]

[a] National Technical University of Athens, School of Electrical and Computer Engineering 9, H. Polytechniou str., Zografou, Athens, 15773, Greece, (tpar, raftop,kntal)@image.ntua.gr

[b] Mount Sinai Foundation, 26, Dorileou str., Athens, 11521, Greece, sinaiticarchiveofmonuments@gmail.com

KEY WORDS: Image classification, Image matching, Feature extraction.

ABSTRACT:

Over the last decade several initiatives were carried out worldwide, towards preservation and promotion of cultural heritage. In this framework, many Byzantine artworks of historic, cultural and artistic interest have been digitized, producing large volumes of visual data, and thus motivating the need for effective indexing and retrieval from such repositories. This paper proposes a decision making framework for content-based retrieval of Byzantine icons, based on a combination of low-level features. In particular initially characteristic icons are selected and a features model is built for the entity under consideration. Next head areas are unsupervisedly detected using Hough transform and the special metrics system of Byzantine art, described by Dionysios from Fourna. Next for each head area proper features, such as MPEG-7 descriptors for shape and color, are extracted. A comparison is then performed between four standard classification algorithms before the decision is made in favour of the multilayer-perceptron with back propagation learning (NNBP). The analysis and results of this classifier indicate the effectiveness, generality, and flexibility of the proposed scheme.

1. INTRODUCTION

The term "Byzantine art" is commonly used to describe the artistic production of Byzantine Empire, from about the 4th until the 15th century. Painters of Byzantine artworks use specific iconographic patterns for the creation of figures. For each pattern, there appear also variations, due to the evolution of the artistic tendencies. Patterns and their variations are quite clearly specified in theoretical treatises followed by painters. The "Interpretation of the Byzantine Art", written in the 18th century by Dionysios from Fourna, is one of the few treatises that have survived until now (Hetherington P., 1974).

The study of byzantine works of art mainly focuses on the recognition and classification of the patterns and variations in each figure. This procedure leads to the chronology of the works of art and their attribution to the artistic ambience in which they were created.

On the other hand over the last decade several governments, organizations, museums and content owners have focused on the preservation and promotion of cultural heritage. In this framework large artistic collections have been digitized, among them a great number of Byzantine works. Obviously artistic repositories, similarly to any other repository, need effective indexing and retrieval tools and techniques and towards this direction a number of methods have been proposed. In particular in Kushki A. et al, 2004, a unified framework for similarity calculation is proposed for the combination of low-level feature similarities in paintings and engravings. The problem of similarity calculation is reformulated as a problem of decision making, where partial decisions are aggregated to achieve overall similarity.

Learning-based characterization of fine art painting styles is proposed in (Li J., and Wang J. Z., 2004.).The authors focus on comparing the painting styles of artists using a mixture of 2-D multi resolution HMMs. The technique is applied on Chinese ink paintings. In (Chan et al 2001), the problem of content-based retrieval of paintings is addressed as a sub-region matching problem, while in (Corridoni J. M., Del Bimbo A., and Pala P., 1999) the PICASSO CBIR system is presented, which supports semantic queries on contrast and harmony based on inter-region color relationships.

In (Lay J. A. and Guan L., 2004) a concept-oriented retrieval engine is proposed, based on the generative grammar of elemental concepts methodology. In this case the language by which color artistry concepts are communicated in artworks is used to operate semantic searches. The color artistry language is explicated into elemental concepts and the associated generative grammar. The elemental concepts are used to index the artworks, while the generative grammar is used to facilitate post-coordinate expression of color artistry concept queries by using the elemental concepts.

In (Lewis P. H., et al, 2004) a retrieval system, applicable to museum and gallery image collections, is presented. Specialist algorithms, such as the multi-scale color coherence vector (MCCV) technique that can provide effective sub-image retrieval even in different resolutions, are combined with conventional content and metadata retrieval approaches to provide cross-collection searching and navigation. Furthermore there are also several works relevant to modelling, intelligent retrieval and presentation of cultural heritage objects and images (Godin G., et al, 2002, Miyazaki D., et al, 2000, Drigas A.S., et al, 2006, Bernardini F.,et al, 2002)

However most of the aforementioned techniques do not straightforwardly consider modelling or retrieving of Byzantine art. Thus they cannot be effectively applied to this domain without considering conjectural styles and incorporating domain knowledge and rules. Furthermore most of the existing techniques relevant to Byzantine art focus on detection of artefacts and image restoration. Characteristic methodologies of these approaches are presented in (, Sotiropoulou S., et al, 2000, and Nikolaidis N. and Pitas I., 2001).

In order to bridge these gaps in this paper we aim at developing a novel CBIR system for Byzantine icons, based on the detection and classification of face regions. Here we focus on the figures of Jesus Christ and Virgin Mary since they appear in several digitized works of two great painters, Emmanuel Panselinos (14th c.) and Theophanes the Cretan (16th c.). Additionally they have been analytically described in the theoretical approach of Dionysios from Fourna, an approach that offers fundamental knowledge and the essential rules for analyzing and interpreting Byzantine artworks.

In the following sections the proposed system is presented in detail. In particular in section 2 the metric rules of Byzantine iconography, as outlined by Dionysios from Fourna, are briefly described. In section 3 the special head area detection module is presented, while the indexing and retrieval methods are explicated analyzed in section 4. Experimental results can be found in section 5, while section 6 concludes this paper.

2. THE METRIC RULES OF BYZANTINE ICONOGRAPHY

In our approach, we take into consideration both similarities and differences between the rules used by Emmanuel Panselinos and those used by Theophanes the Cretan – cited by Dionysios from Fourna.

Painters of the Byzantine works of art follow the specific instructions that Dionysios from Fourna had recorded for painting holy figures. According to these instructions, initially a painter separates the painting area into seven semantic segments of equal size, each of which has specific characteristics that can make the figure distinguishable. An example is presented in Figure 1, where the standing holy figure of Jesus Christ is presented.

Let us denote by P the height of each of these segments. Starting from top to down, the first segment (no. 1) is separated into 4 equal smaller parts. If H is the height of each of these parts, then

$$H = \frac{P}{4}, \qquad (1)$$

The first, second, third and fourth smaller parts should contain the hair, forehead, nose and the area from mouth to chin respectively, according to the specifications of Dionysios from Fourna. These four parts, which compose the first segment, make up the head of the holy figure.

The second segment, also of height P, contains the part from neck to thorax while the third segment contains the part from thorax to elbow and waist which always lay at the same height. Next, in the fourth segment the abdominal area is usually depicted while the fifth segment contains the area from legs until the knees. The sixth segment contains from ankle to foot and finally the feet of the figure are located at the seventh segment.

Additionally there are 4 general types of icons in byzantine iconography of which the smaller contains a figure of height $3 \cdot P$ and the larger a figure of height $7 \cdot P$ ($3 \cdot P < H < 7 \cdot P$). Thus the area A, in which a figure is contained, follows the inequality:

$$3 \cdot P \leq A \leq 7 \cdot P \Rightarrow 12 \cdot H \leq A \leq 28 \cdot H \qquad (2)$$

Furthermore, usually the height Y of the whole image is $Y = A + 2 \cdot H$, thus

$$14 \cdot H \leq Y \leq 30 \cdot H \Rightarrow Y/30 \leq H \leq Y/14 \qquad (3)$$

The last inequality relates the height H of the small parts to the height Y of the whole image. According to Dionysios from Fourna, the head of a saint should be surrounded by a halo, a circle that signifies the Holy Spirit (see Figure 1). In order to paint the halo, the painter draws a circle, centered at the middle point of the nose with a radius:

$$R = 2.5 \cdot H \qquad (4)$$

Figure 1: Metric rules of Byzantine iconography

Based on (3) and (4) the head area can be detected by searching for circles of radius around R as described in the following section.

3. HEAD AREA DETECTION

Digitized Byzantine artworks present many challenging issues regarding low level image analysis. In this section, we describe our method for extracting the head areas from saints' figures in various such images (Figure 2), for the purpose of classification of saint figure.

Our key observation for extracting the head area of a saint that is depicted in a picture is by identifying the halo around the saint's head, which, according to the treatise of Byzantine iconography, is always a circle and thus even in noisy images, like the ones we are dealing with, can be efficiently detected by the generalized Hough transform (John G. and Langley P.,1985) for circle detection. Looking for the halo circles in the picture serves also the purpose of identifying only the saints in the image and distinguishes them from other figures that may coexist in the same image. At the same time the diameter of the halo, once identified, determines the scale of the image thus permitting the use of a broader range of MPEG-7 descriptors that are not necessarily invariant to scaling, for the purpose of classification.

Figure 2: Holy figures

Once the halo(s) have been identified a multiple threshold technique can be applied by calculating the average image intensity values in homocentric circles inscribed into the main halo circle. This way the region that contains the head can be identified and most of the area between the head and the halo can be removed. Finally, by applying suitable median filters the heads can be extracted with very good accuracy.

For the halo identification, we use the Hough transform, one of the most popular voting methods for parameter estimation. In the Hough transform, each point on an edge votes for several combinations of parameters. The parameters that win the majority of vote are called the winners. The approach for fitting a circle to data can be considered as following:

1. Quantization of the parameter space with regard to the parameters a and b of the Hough transform.
2. Assign an accumulator to each cell in the parameter space and initialize all accumulators $M(a,b)$ to zero.
3. Compute the gradient direction $\theta(x,y)$ and magnitude $G(x,y)$ for all the edge points in the image.
4. For each edge point $G(x,y)$ increment all points in the accumulator array $M(a,b)$ along the line: $b = a \cdot \tan\theta - x \cdot \tan\theta + y$.
5. Find the local maxima in the accumulator array and determine the center of the circles.

The algorithm above does not assume any knowledge about the halo location or its radius therefore it can be optimized if we take under consideration the pertinent rules of Byzantine iconography regarding the halo topology and radius that have been presented in the previous section.

After the halo circle has been identified, we estimate two thresholds by drawing two concentric circles inscribed into the halo. From the small circle we estimate the average intensity value for the actual head and from the ring between the greater circle and the halo we estimate the intensity value for the halo. By choosing a threshold between these two values, we segment the area in the ring between the two circles in two regions, head and halo.

Finally, we apply a median filter with appropriate size to the segmented image in order to produce masks that better isolate the extracted head area. Application of these masks to the original images produces the final images of extracted heads that will be used in the next section for classification. The pictures in Figure 3(a) show, from left to right first, the stages for the head extraction for one image (extraction from halo location, threshold calculation, median filtering, and final extraction) and in Figure 3(b) the extracted heads for nine such images are shown.

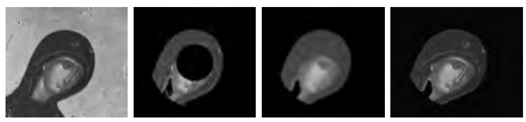


Figure 3(a): Head extraction for one Holy figure

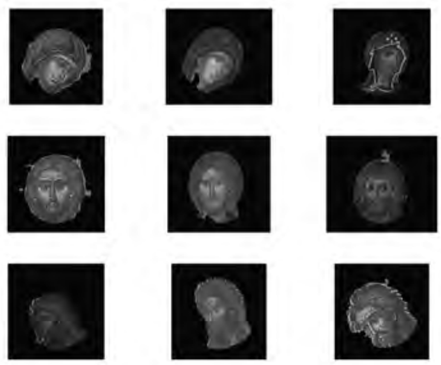


Figure 3(b): Extracted heads of several Holy figures

4. CLASSIFICATION

The problem of classification is consisted in learning a target function that assigns unknown objects into predetermined categories. The purpose of the learning procedure is the creation of a correspondence between distinctive values and unknown objects. The collection of values from representative samples from each category is a prerequisite for problem solution because these samples have been classified from experts and constitute the training data set.

A machine learning algorithm receives as input the samples of the training data set each of which is represented by a feature vector in order to extract the knowledge needed by the target function. According to the type of information a feature can be continuous (e.g. a real number), nominal or a set of distinguished values. In this way each data element is considered as a point in a n-dimensional space.

In this section we evaluate the efficiency of the standard classifiers in the framework of Saint classification. The algorithms that will be compared are a) Naive Bayes (NB) and b) Flexible Naïve Bayes (FNB) (Rumelhart, D. E., et al, 1986.), c) Support Vector Machines (SVM) 0 (Vapnik V., et al, 2006) and d)multilayer-perceptron with back propagation learning (NNBP).

The input to these classifiers will be a vector representation of the extracted head areas. To evaluate the accuracy of classification across testing, we use the *F-score* which is a combination of the precision and recall.

$$F - score = \frac{2 \cdot precision \cdot recall}{precision + recall} \qquad (5)$$

5. CLASSIFICATION EXPERIMENTS

In this framework, we examine the performance of the standard classifiers, in the case of Saint recognition in Byzantine art. The Naive Bayes, Flexible Naïve Bayes, multilayer NN with back propagation and the SVM have been selected and their efficiency is estimated in the framework of Byzantine icons.

The dataset for our experiment is made of 2000 byzantine images which belong to three classes: *Jesus, Virgin Mary and other Saint.* In Figure 4 five such images from each class are shown. Initially, the dataset is divided in a random fashion into two parts, keeping the class proportions equal to the ones presented in the whole dataset.

The first part is consisted of 800 byzantine icons and is considered the training set D, while the rest (1200 byzantine icons) compose the testing set UD. The dataset D is further separated into two equal subsets, PD with the positive and ND with the negative samples.

Our experiments focus on the correct classification of Jesus Christ and Virgin Mary as opposed to other Saints.

For this purpose, we use the MPEG-7 descriptors (Chang S., Sikora T. and Puri A., 2000) and (Manjunath B. S., Salembier P. and Sikora T. 2001) of Scalable Color, Color Structure, Edge Histogram and Region Shape as well as Hu moments, (Hu M. K., 1962),
to represent each extracted head area. To reduce the dimensionality of the correlated MPEG-7 vector descriptors, we use principal component analysis to eliminate those components that contribute less than 2% to the total variation in the data set. The resulting data set is uncorrelated with zero mean and variance equal to 1.

The four classification algorithms, NB, FNB, SVM and NNBP are trained using subsets PD and ND and the four respective classifiers are created, one for each.

Figure 4: Extracted head areas of Jesus (1), Virgin Mary (2) and other Saints (3) of training dataset: Positive samples from classes of the training dataset.

The generalization capability of the classifiers is then measured against the subset UD using the F-score measure. The whole procedure is repeated 50 times in order to reduce statistical measure variations.

In Figure 4, five positive samples of the training dataset for each class are shown. The head areas of the training data set after extracted from the initial images are then represented by four MPEG7 descriptors and Hu moment vector representation. The dimensionality of the MPEG7 input vectors was reduced to 10,10,9 and 7 respectively as a result of the principal component analysis treatment.

In Table 5 the mean value and standard variation of F-score for each algorithm are depicted. The NNBP classifier for the specific subdataset UD provides better classification than the other classifiers as is depicted in Table I.

Table I: Comparison of Classification algorithms

	NB	FNB	SVM	NNBP
Mean Value	0.347	0.541	0.608	0.895
Standard deviation	0.020	0.012	0.086	0.038

For the NNBP in particular the four MPEG7 and Hu moments vector representations were used to train five neural networks each one consisting of two hidden layers and varying the number of neurons in the first layer between 3 and 7 and the number of neurons in the second hidden layer from 1 to 3. In all cases the networks were able to learn the correct classification for all the presented images of the training set.

The NN generalization capability was then improved by regularization under the Bayesian framework of (John G. and Langley P.,1985) according to which the weights and biases of the network are assumed to be random variables with specified distributions and the regularization parameters are related to the unknown variances associated with these distributions. We can then estimate these parameters using statistical techniques. A detailed discussion of Bayesian regularization can be found in (MacKay D. J. C., 1992, and Foresee F. D. and Hagan M. T., 1997). The final classification decision was taken after voting on the outputs of the 5 NN classifiers and correct recognition was achieved at 90% of these cases, which renders the technique successful.

In Figure 6a samples of corrected classification results of testing dataset are depicted. There are three classes: 1. Jesus, 2. Virgin Mary and 3. Other Saints. For all classes, a random set of 10 corrected classified samples are depicted.

On the contrary, Figure 6b shows some missed classification results and the wrong class number that was assigned to these pictures is shown in the first column. As we can see in the first row, heads that look very similar to Jesus, like Saint Artemios (first line, fourth column) are misclassified in the first class.
Males with no beard, partial heads results of poor head extraction or females with kerchief similar to Virgin Mary's are misclassified in the second class, as we can see in row 2.
Heads of Jesus and Virgin Mary are misclassified in the third class due to low resolution, blurring, and other noise effects as is shown in the third row.

By inspecting the cases of misclassification we can identify some factors that seem to play important role in the accuracy of the proposed algorithm:

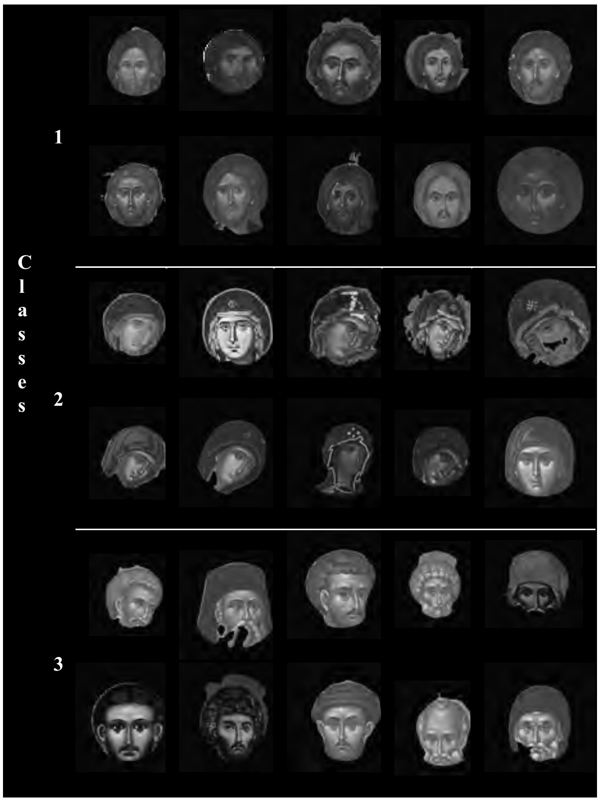
- The almost identical similarity of some Saint Heads presented in Holy images suggests the use of other picture contents like clothes, artifacts, surrounding objects that need to be taken under consideration for distinguishing Saints with identical heads.
- The noise affects the classification accuracy.
- The accuracy of the head extraction algorithm is a critical factor towards correct classification.

6. CONCLUSION

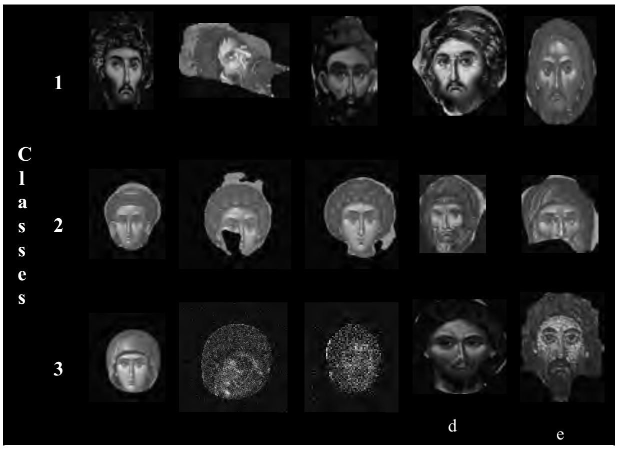
In this paper a CBIR system for Byzantine artworks has been proposed. The system incorporates the metric rules of Dionysios from Fourna, Hough transform for head area detection and NN classifiers for entity determination. A merit of the proposed system is that it can support artists, Byzantine icon analysts, historians, archaeologists and researchers better analyze, interpret and finally understand Byzantine works of art, their correlations, common characteristics, color and style evolution through centuries and many other features.
Currently these processes are traditionally based on the good observation and perception capabilities of historians of art. Thus research towards pattern analysis and classification of regions in Byzantine icons will positively contribute to the elaboration of expert systems for mining and recognition of figures and variations of figures of Byzantine art.

ACKNOWLEDGEMENT

This research is performed in the framework of the Greek Secretariat of Research and Technology Project "MORFES", which is co-funded by the European Social Fund (75%) and National Resources (25%).



a) Correct classification



b) Missed classification

Figure 6: Samples of classification results

REFERENCES

Bernardini F., Rushmeier H., Martin I. M., Mittleman J., and Taubin G., 2002, Building a digital model of Michelangelo's Florentine Pieta, *IEEE Computer Graphics and Applicaions*, 22(1), pp. 59–67.

Chan S., Martinez K., Lewis P., Lahanier C., and Stevenson J., 2001. Handling sub-image queries in content-based retrieval of high resolution art images, Processing on. International Cultural Heritage Informatics Meeting, vol. 2, pp. 157–163.

Chang S., Sikora T. and Puri A., 2000, *Overview of the MPEG-7 standard.*

Corridoni J. M., Del Bimbo A., and Pala P., 1999. Retrieval in paintings using effects induced by color features, *IEEE Transactions on Multimedia*, 6(3), pp. 38–53

Drigas A.S., Koukianakis L.G., Glentzes J.G., 2006, A dynamical system for e-culture services, *WSEAS Transactions on Systems*, 1(5), pp. 125-130.

Foresee F. D. and Hagan M. T., 1997, Gauss-Newton approximation to Bayesian regularization, in *Proceedings of International Joint Conference on Neural Networks*, pp. 1930–1935.

Godin G., Beraldin J. A., Taylor J., Cournoyer L., Rioux M., El-Hackim S., Baribeau R., Blais F., Boulanger P., Domey J., and Picard M., 2002. Active optical 3-D imaging for heritage applications, *IEEE Computer Graphics and Applications*., pp. 24–36

Hu M. K., 1962, Visual Pattern Recognition by Moment Invariants, *IRE Trans. Info. Theory*, IT-8, pp.179-187.

Hetherington P., 1974. *The "Painter's Manual" of Dionysius of Fourna,* Sagittarius Press, London.

John G. and Langley P.,1985. Uncertainty in Artificial Intelligence, *Morgan Kauffman, San Mateo*

Kushki A., Androutsos P., Plataniotis K. N., and Venetsanopoulos A. N., 2004. Retrieval of Images From Artistic Repositories Using a Decision Fusion Framework, *IEEE Transactions on Image Processing*, 13(3).

Li J., and Wang J. Z., 2004. Studying Digital Imagery of Ancient Paintings by Mixtures of Stochastic Models *IEEE Transactions on Image Processing*, 13(3).

Lay J. A. and Guan L., 2004. Retrieval For Color Artistry Concepts, *IEEE Transactions on Image Processing*, 13(3).

Lewis P. H., Martinez K., Abas F. S., Fauzi M. F. A., Chan S. C. Y., Addis M. J., Boniface M. J., Grimwood P., Stevenson A., Lahanier C., and Stevenson J., 2004, An Integrated Content and Metadata Based Retrieval System for Art, *IEEE Transactions on Image Processing*, 13(3).

MacKay D. J. C., 1992, Bayesian interpolation, *Neural Computation*, 4(3), pp. 415–447

Manjunath B. S., Salembier P. and Sikora T. 2001, *Introduction to MPEG-7*, JOHN WILEY & SONS, LTD.

Miyazaki D., Ooishian T., Nishikawa T., Sagawa R., Nishino K., Tomomatsu T., Takase Y., and K. Ikeuchi, 2000, The great buddha project: modelling cultural heritage through observation, *in Processing. of International Conference Virtual Systems and MultiMedia* (VSMM 2000), pp. 138–145.

Nikolaidis N. and Pitas I., 2001, Digital image processing in painting restoration and archiving, *IEEE International Conference on Image Processing.*

Ballard D. H., 1981, Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recognition*, 13(2), pp. 111–122.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1986, *Learning internal representations by error propagation*", Parallel Distributed Processing: Explorations in the Microstructure of Cognition.Foundations, MIT Press, Cambridge, MA.

Sotiropoulou S., Karagiannis G. and Chryssoulakis Y., 2000, Colour Correction and Storage of Digitized Image Captures Representing Art Paintings, *in 1st International Conference on Color in Graphics and Image Processing*, Saint-Etienne , France.

Vapnik V., and Kotz S.,2006. *Estimation of Dependences Based on Empirical Data,* Springer *[this is a reprint of Vapnik's early book describing philosophy behind SVM approach. The 2006 Appendix describes recent developments].*

Virtual Reality Applications in Cultural Heritage

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

WORTH A THOUSAND WORDS? THE USEFULNESS OF IMMERSIVE VIRTUAL REALITY FOR LEARNING IN CULTURAL HERITAGE SETTINGS

Laia Pujol Tost [a], Maria Economou [b*]

[a] Centre for Museology, School of Arts, Histories and Cultures, The University of Manchester, Oxford Road, Manchester M13 9PL, UK - laia.pujol@uab.cat
[b] Museology Laboratory, Dept. of Cultural Technology & Communication, University of the Aegean, Har. Trikoupi & Faonos Str, Mytilene 81100, Greece - m.economou@aegean.gr

KEY WORDS: Virtual Reality, Cultural Heritage, learning, ICT perception, evaluation, visitor studies, design guidelines, Hellenic Cosmos, Foundation of the Hellenic World.

ABSTRACT:

The goal of this paper is to investigate whether immersive Virtual Reality is suitable for learning about archaeology and the past in Cultural Heritage settings. To that end it presents the conclusions related to learning from the visitors' survey undertaken in 2007 by the Museology Laboratory of the University of the Aegean at Hellenic Cosmos (the exhibition centre of the Foundation of the Hellenic World in Athens) and contrasts these with other similar studies. This project was aimed at comparing the learning outcomes, perception and use by audiences of two different VR systems and a related exhibition. It included qualitative and quantitative analysis of the data gathered through in situ observations, interviews with museum education guides and face-to-face questionnaires with visitors. The results confirmed that, as previous studies have shown, VR systems allow a different kind of learning, but also questioned the common believe about their advantage for children in comparison with other interpretation methods.

1. INTRODUCTION

The introduction of Virtual Reality (VR) in the Cultural Heritage (CH) field as both a communication and a learning tool over the last decade has attracted considerable interest from researchers and CH practitioners in evaluating its effectiveness in both directions. Audiences and professionals within the sector –archaeologists, cultural mediators, ICT specialists– commonly believe that VR is especially suitable for children and for informal learning environments, such as CH settings. This belief is based on some VR features shared with traditional means of museum communication and on the extrapolation of previous results obtained in formal learning environments with regard to the most suitable subjects, cognitive gains and learning processes when using these applications.

However, research has shown that Virtual Learning Environments (or VLE, as pedagogical computer applications are called in the formal learning environment) cannot be considered in the same way as usual museographical resources because even when transferred to an exhibition, their design as exhibits is determined by their computational interface (vom Lehn, Heath et al. 2005). Similarly, informal learning environments, such as exhibitions, are different from formal learning environments with regard to the goals, availability of users' time for exploring the resources, or the role of social interaction. Therefore, there is a need to evaluate in situ the effectiveness of ICT applications for learning in CH settings, taking into account the specific goals and characteristics of these contexts.

The goal of this paper is to investigate whether immersive VR is suitable for learning about archaeology and the past in CH settings. To that end it will present the results related to learning from the visitors' survey undertaken in 2007 by the Museology Laboratory of the University of the Aegean at the Hellenic Cosmos, the cultural centre of the Foundation of the Hellenic World in Athens, and compare these with previous studies in order to reach some conclusions about the potential usefulness of VR as a learning tool for CH.

2. THE USEFULNESS OF VR FOR LEARNING IN CH SETTINGS

In spite of the high cost of VR devices, museums have been introducing them within their usual repertoire of communication tools for three main reasons (Roussou, 2000): commercial pressures; the capacity to transcend the physical space of the exhibition with its potential for virtual reconstructions; and its ability to support the educational role of the museum. Only few studies have exclusively focused on ICT's effectiveness for learning in CH settings. Until now this aspect was more often included within wider visitor surveys or was extrapolated from the formal learning environments, where the effectiveness of VLE has been systematically studied. The related literature (see (Economou and Pujol, 2006) for a critical review) shows that, especially in the later contexts, technological applications have positive results with regard to impact, iconic skills and improvement of engagement. However, in the case of exhibitions, some negative aspects have also been evidenced, such as problems of integration with the rest of the exhibits, obstruction of learning by non-intuitive interfaces, and limitation of social interaction due to the one-to-one communication paradigm of these applications.

One of the oldest specific examples in the CH field is a study (Pimentel and Teixeira, 1995) aimed at evaluating the potential of VR for education and emotional satisfaction in a simulation of the real world at the Computer Museum in Boston. The observations inside the room showed that people spontaneously preferred to learn through the interactive experience instead of reading text, which was systematically marginalised. The main problem in that application was to get familiar with the interface: visitors spent between 5-10 minutes to understand the rules and aims of the simulation.

The survey conducted at the temporary exhibition "Immaginare Roma Antica" held in Rome in 2005 (Forte, Pescarin et al., 2006) analyzed the perception and use of different kinds of ICT exhibits. The factors attributed by visitors to the suitability for learning were: richness of information (in the case of multimedia); quality of reconstruction (in the case of VR applications); and in all cases, link with previous knowledge. Nevertheless, observations showed that these factors would have no effect on learning if the interface and the organization of the content were not intuitive.

A study analyzing the effects of immersion in several science museums used a typology based on the kind of representation and knowledge domain (Belaën, 2003). The results showed that immersive devices do not guarantee an immediate acquisition of knowledge or positive attitude towards the experience (with the exception of younger users) and that the operation needs to be as multi-sensorial and natural as possible because, otherwise the understanding of the content is compromised.

Finally, the contribution of computational interactivity was analyzed through an in-depth evaluation at the Ename Archaeological Museum in Belgium in 2006 (Pujol and Economou, 2007). According to visitors' opinion, technology is appropriate for learning because it allows a flexible, personalized exploration of a richer quantity of information. In the case of VR, it offers the possibility to reconstruct and manipulate elements (buildings, objects) or phenomena (historical processes) which cannot be seen anymore. Nevertheless, researchers also understood that in the last case, images need to be supported by a verbal discourse because since they tend to be photorealistic, they cannot represent abstract elements such as external causal agents, social relationships, etc., where more schematic representations and especially verbal language excel. Nevertheless, observations evidenced again that there can only be understanding/learning when all problems related with the interface design are previously solved.

The conclusion from these studies is that VR is attractive and motivational, especially younger for visitors, and perceived as a useful tool for learning descriptive content about objects and processes. Nevertheless, learning or understanding will only be achieved if the physical and virtual interfaces are multisensorial and intuitive enough.

3. EVALUATION AT THE HELLENIC COSMOS, FOUNDATION OF THE HELLENIC WORLD

3.1. The Hellenic Cosmos

The Foundation of the Hellenic World (Gaitatzes, Christopoulos et al., 2000) is a Greek privately funded non-profit CH institution based in Athens, Greece. It aims at the preservation and dissemination of Hellenic history and culture in order to create an awareness of the universal dimension of Hellenism and its active contribution to cultural evolution. VR technology is used in two ways at the Foundation: as a tool for research and as a tool to disseminate this knowledge. The latter is mainly done through the Hellenic Cosmos (figure 1), a cultural centre founded in 1998 and based in a former industrial area of Athens now in transformation, which uses state-of-the-art ICT and museological trends to offer educational programmes, virtual reconstructions, documentaries and exhibitions.

Figure 1: The Hellenic Cosmos in Athens

At the time the survey took place, the Hellenic Cosmos had three main exhibits involving ICT on display. Kivotos, inaugurated in 1999, is the oldest system. It is a CAVE-like system composed by 4 back-projection screens delimiting an area of $3m^3$, where a group of maximum 10 people is immersed in different shows, mainly reconstructions of archaeological sites, such as Olympia or the ancient city of Miletus. Users wear stereoscopic glasses, which create a real 3D effect. The interaction with the system is controlled by the museum educator, who moves within the virtual world with the help of a wand and a tracking system on her head.

Tholos, the newest system, inaugurated in December 2006, is a virtual reality theatre made of a semi-spherical screen inclined 23º and surrounding a room with 132 seats, each one equipped with a joystick and four buttons. The evaluated application, the first one designed for this system, presents a reconstruction of the ancient agora of Athens in three different periods. The interaction with the system takes again the form of a spatial navigation and is led by a first museum educator with the help of a wand, while a second guide provides the archaeological content (description of the buildings' functions). Finally, the exhibition, complementing Tholos' content, presents information about the evolution of the agora, its historical meaning and institutions, and how these have been transmitted until today. It combines different types of exhibits (text, images, hands-on and technological applications), which allow different kinds of exploration and degrees of collaborative interaction.

The unique concentration of VR systems and high-tech exhibits, combining different degrees of interactivity, immersivity and social interaction, made this cultural centre very suitable as a case study for our research investigating the suitability of ICT for CH settings and its perception and use by audiences. At the same time, it allowed us to refine our methodology specifically designed for the evaluation of technological exhibits in museums, which takes into account and ultimately integrates all the factors involved in their use: the interface features, the visitors' characteristics and the context (spatial configuration and social interactions).

3.2. Goals and methodology

The visitor survey conducted at the Hellenic Cosmos of the Foundation of the Hellenic World had two specific goals. The first was to investigate the added value of Presence (see (Schuemie, Van der Straaten et al., 2001) for a comprehensive definition of the term) in the CH field. The second goal of the survey had to do with the suitability of high-tech exhibits for learning in CH museums. Given the results presented in the previous section, we were interested in analysing visitors' preferences about communication solutions for learning, and also tried to understand (using short-term recall and in-situ observations) what remained in their mind from these diverse experiences.

To meet these goals we used different methods (interviews with visitors, staff and in situ observation) in order to record different kinds of information. By combining the advantages of qualitative methods (interpretation of answers and behaviours) and quantitative methods (statistical analysis of the previous) we expected to gain a better understanding of the subject by obtaining new knowledge and establishing relationships and explanative hypotheses. With regard to observation, the pilot tests demonstrated that, due to the specific settings, observation inside the Tholos and the Kivotos provided very poor or already known results. Consequently, we decided to concentrate it on the exhibition, where a standardized sheet with a floor plan and behaviour codes were used in order to record visitors' timings at each exhibit, paths and behaviours. These data would allow a comparative analysis of the use of high- and low-tech, interactive and "traditional" exhibits in order to detect and understand preferences and social uses.

The role and opinion of museum educators played an important role in this research project, since they are the main navigators of the VR applications and mediators in the construction of meanings at the Hellenic Cosmos. Therefore, we conducted semi-structured recorded interviews with the guides, who told us about the usefulness of high- and low-tech exhibits; the role of guides; and visitors' reactions and behaviours. However, the most important source of information was the semi-structured interviews with visitors. The questionnaires, both in Greek and in English, combined open-ended, likert and categorized questions, distributed in the following five sections: general questions about the HK (most liked and disliked exhibits); comparative questions about the feeling of Presence; preferences in communication solutions for learning (usefulness of the exhibits, suitability for learning, best way to convey information, content recall); quick cross questions about the exhibition; and finally, demographic data.

After pilot testing of the questionnaire and observations in spring 2007, interviews with visitors were conducted for 15 days in July 2007. The survey was carried out by 4 interviewers, who collected a total amount of 97 interviews, 75 in Greek with Greek visitors and 22 in English, mainly with tourists. Interviews lasted between 15 and 30 minutes. After the data gathering, we carried out a qualitative as well as statistical analysis (categorization of open-ended answers, univariate and bivariate techniques), the results of which are presented below.

3.3. Results

The third part of the questionnaire was devoted to learning. The first question (table 1) asked about the usefulness of the VR applications they had seen. As all VR exhibits were included in this overall count, the answers provide a general idea about visitors' perception of the Hellenic Cosmos: it allows learning about places in the past thanks to the visual and experiential component of its exhibits.

Usefulness of seen applications	Freq.	%
Better than books because experiential/ enjoyable/visual/strongest impact	17	12,9
Feeling of transportation/participation/ experience of the past (life, buildings)	12	9,1
Useful for learning (provides new information in a little time and nice way)	12	9,1
Provides a reconstruction/image of buildings/the place (in the past)	10	7,6
Visual component facilitates understanding/fixation/attention/accessibility	10	7,6

Table 1: Usefulness of seen exhibitions

With regard to the specific contribution of each exhibit (figure 2), we found out that the Tholos was mainly appreciated by its experiential character. In the case of the Kivotos, the "feeling of transportation / participation / experience of the past (people, life, buildings)" was the main contribution. With regard to the exhibition, its technological displays were also considered better than books thanks to their experiential and participative component. This question evidences the importance of the experiential component for learning (in comparison to books), which can be therefore considered the main characteristic of the Hellenic Cosmos as an informal learning environment.

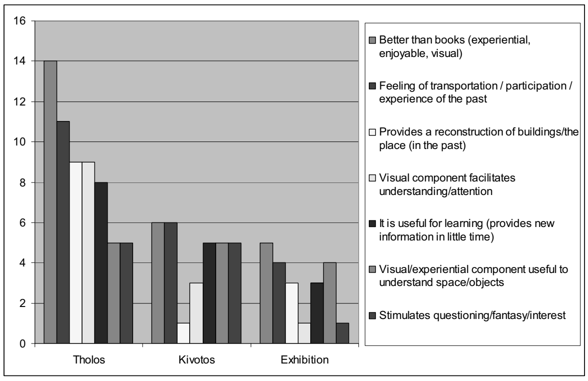


Figure 2: Specific contribution of each exhibit.

Once we had obtained a general idea of how visitors perceived VR applications, we wanted to understand how they related them with a specific exhibition. The Hellenic Cosmos provides a good example because the Tholos and the exhibition present the same subject and Kivotos' archaeological reconstructions can be considered quite similar to the 3D reconstruction of the "Ancient Agora". In case visitors had not visited the exhibition, we asked them to imagine an exhibition with this subject and we took it into account in the contingency table. Both categories of visitors answered that VR exhibits provided an introduction, a general image about spatial aspects; while the exhibition, because it contained text, photographs and objects, provided more or more "substantial" (meaning factual or historical) information. This difference in the learning outcomes was independently confirmed, as we will see later, thanks to the question about content recall.

Given these results, we asked visitors which was the most suitable exhibit for learning about ancient Greece (table 2). The Tholos and the exhibition were most frequently mentioned, but taking into account the relative number of interviewed visitors in each one (86 for Tholos and 23 for the exhibition), it is evident that the exhibition was considered by far the most suitable exhibit for learning. This is confirmed by the general questions about the Hellenic Cosmos. In this case, learning was selected by visitors as the second most important reason for choosing the exhibition as the most liked exhibit, but it was marginal for the VR systems. The reasons given were that that the exhibition provided "more detailed / well documented knowledge/explanations" and especially "new / more information about people / history / democracy" (10/19); was "richer in communication means" (4/19); and allowed visitors to "have control over their visit (time, activities)" (4/19).

Exhibit	Freq.	%
None	1	1,9
Tholos	19	35,2
Kivotos	5	9,3
Exhibition	19	35,2
All	2	3,7
Tholos + Kivotos	4	7,4
Tholos + Exhibition	4	7,4
Total	54	100
System	61	
Total	115	

Table 2: Best exhibit for learning about ancient Greece.

These results do not exclude the usefulness of the technological exhibits, which were more suitable to "perceive/describe the remains/place/spatial details" (4/18 for the Tholos and 4/5 for the Kivotos). More specifically, according to visitors, the Tholos had the advantage that it "combines images and oral explanations" (3/18) and therefore "provides a good general idea within limited time and place" which made it "good or easier for wide audience" and especially for children. The other VR system, the Kivotos, was also considered "useful to perceive/describe the remains/place/spatial details" (4/5) but in this case, its main advantages were the possibility to establish a closer relationship with the exhibit and the guide, as well as to play a more active role (2/5).

Communicational preferences	Freq.	%
Human guide	74	68,5
Audio commentary	11	10,2
Audio-guide	11	10,2
Images alone	8	7,4
Text	4	3,7
Total	108	100

Table 3: Communicational preferences for learning.

The next question was related to visitors' preferences in communication preferences for learning (table 3). Visitors considered that the best way to explore the application at both VR applications was with a human guide, because this allows a direct interpersonal interaction, thanks to which he/she immediately adapts to the audience and can solve doubts or provide more information (20/63). Another reason was that some visitors do not like machines and think that "a human is more immediate/alive/spontaneous/non standardized" (20/63). The exception to this was children, who although recognized that guides can solve doubts and give more information, they still preferred to explore the virtual world alone because they like the technology. With regard to the rest of visitors' preferences we found out that audio-guides were preferred because they "allow choosing language/level of information/control over your visit" (3/4) or by those who "did not like the guide (explanation/articulation/attitude)" (2/7). Those who wanted the audio commentary (4/7) had had problems with the guide ("voice/navigation style disrupts the feeling of being there"). Text was chosen because it was considered "better for learning" (2/3) and could be taken away (1/3). Finally, images alone were chosen because of individual skills and preferences by people who "prefer (simple instructions and) individual exploration" (3/3).

Through visitors' answers in this section of the questionnaire and in the general questions about the HK, we found out that the different exhibits (technological vs exhibition) were probably allowing different kinds of learning. By analyzing which specific content visitors remembered from their experience at the different exhibits, either something they did not know before or something that impressed them, we could verify if this was true. In this way we could detect some aspects of cognitive gain (recall), while at the same time covering personal differences in learning / communication abilities, and taking into account the fact that, as it has been previously demonstrated (Osberg, 1997; Reid, Zhang et al., 2003), learning at VR displays cannot be measured trough traditional text-based tests because it activates different mental skills. Therefore, we separated the answers by exhibit and we established two different variables: one gathering the detailed answers (with as many categories as answers in each exhibit) and one with 11 categories, valid for all exhibits and aimed at comparing the different kinds of output arising from each one (figure 3).

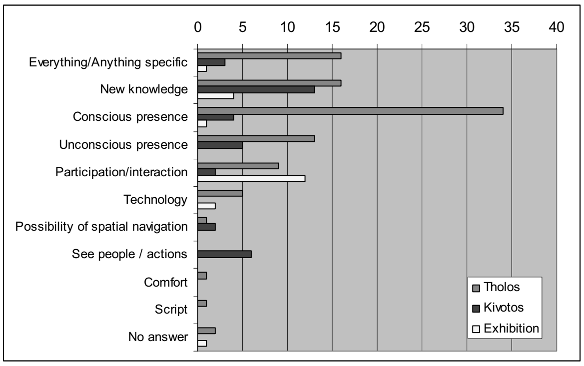


Figure 3: Recall by exhibits (categorized).

The most remembered aspect of the Tholos' content was the transition from Classic to Hellenistic times represented by the building of the Stoa of Attalus. We classified this as conscious presence (34.7%) because, from our own experience, this part of the VR show exerted a very strong impact on our attention due to its combination of powerful (by size and volume) and dynamic audiovisual stimuli. Visitors felt unconscious presence (13.3%, the third most frequent answer) when travelling through the amphora (10 cases), and at some points where they physically felt they were really falling or flying (2 cases). We believe it was stronger at the amphora because the visual field was mainly occupied by one very big element (the ancient vase), moving very quickly, while the sensation of flying involves more visual elements and depends on the guide (smoothness of navigation and coincidence in the point of view). It is worth pointing out that unconscious presence appeared in all age categories because it is an automatic reaction of the human brain, aimed at protecting the person, but conscious presence was comparatively more frequent in younger ages (7-17 year-olds), which is probably related with children's natural capacity to suspend disbelief.

The second most frequent answer was "everything/anything specific" (16.3%). However, another 16.3% of answers reflected there was acquisition of new knowledge, related to elements that were very familiar to visitors' and therefore acted as cognitive connectors. It is interesting to note that it was more difficult for children to say anything specific and to judge they had acquired new knowledge, which means that age is another factor intervening in recording learning outputs. And also gender: a striking 26.41% of women against 6.67% of men gave answers related to knowledge. Finally, in 9.2% of cases, visitors remembered participating (mainly at the ostracism part of the application) and 5.1% were also impressed by the size of the screen (technology).

Despite the fact that in the general questions Kivotos was not associated with learning, we found here that the most remembered thing had to do with the acquisition of new knowledge (37.1%), especially in those above 17 years and in female visitors (66.6% against 36.84% of men). From the interviewees' answers we understood that this was due to the relevance of one of the subjects shown at the Kivotos: the Olympic Games are part of the Greek educational curriculum and of western European identity and therefore, are well known but have many differences with the modern games. The second most frequent answer found across both genders and across all categories of age, was the presence of people performing activities in the virtual applications (17.1%).

The next most frequent answers had to do with Presence but the order was inversed in the Kivotos with respect to the Tholos. In this case, there were more occurrences of unconscious (14.3%) than of conscious presence (11.4%) appearing in all categories of age. In the first case, it was indicated by the reflex movement of Kivotos' participants to avoid the javelin (2 cases), and the corporal sensation of flying (2 cases) and of passing through the objects (1). We believe that the relevant factors here were the true tri-dimensionality provided by the flickering glasses; the higher immersivity of the CAVE-like system (floor screen and isolation from the real world); and the proximity with the guide (bigger coincidence in the point of view). There were also moments of conscious presence when entering the stadium at Olympia (2) and when travelling with the boat (1) or diving at Miletus (1), related to another of the Kivotos' applications on the ancient Greek city in Asia Minor. Here again, acted a combination of dynamic immersive audiovisual stimuli (the noises of the stadium or the sea, the boat's perspective, the immersion inside the tunnel), which demonstrates the importance of capturing the user's attention with strong stimuli that isolate him/her from the real environment.

In the case of the exhibition, the visitors' main recollection, across all interviewed ages, origins and genders, was not related to the acquisition of new knowledge (it was the second, with 19% of appearances) but with the possibility of participation/interaction (57.1%). This is due to the museographic design of the exhibition, which contained multiple interactive exhibits: visitors considered that the exhibition provided new knowledge about the historical dimension of the agora, but were impressed by the multiple hands-on exhibits (9 cases), the big touch screens (4 cases) and the VR multi-user exhibit (2 cases). This answer is confirmed by the specific questions about the exhibition, where interviewees who had visited it, answered that the best exhibits for learning were the hands-on displays (29.2%), the VR multi-user application (20.8%,) and the big touch screens (16.7%), whose virtual schematic/geographic representation, as people emphasized, allowed them to better understand a subject they had been taught at school.

A second reason for the second position of "new knowledge" in visitors' recalls is that due to the structure of the exhibition (made of different units easily identifiable), it was easier for them to mention the exhibit instead of the content: as a result, contents and museographic elements are mixed in our results. An element which would demonstrate this hypothesis is that, unlike the Tholos, at the exhibition only 1 person was unable to mention something specific she had learned. On the other hand, it is interesting to note that here again tourists learnt more than Greeks and that female visitors (33.3%) gave more answers related to knowledge than male ones (13.5%).

3.4. Discussion

One of the main goals of this survey was to investigate audiences' perception of ICT, and particularly VR applications. As we had seen in previous evaluations (Forte, Pescarin et al., 2006; Pujol and Economou, 2007), apart from an eventual interest in technology directly linked with the visitors' professional field (e.g. computing, architecture, etc), ICT are mainly understood as a tool to learn about CH. Nevertheless, subsequent results about the perceived or recorded learning benefits showed that there is a contradiction between users' demands from ICT to support learning and their real use or effectiveness. In the first case, when justifying their choice of the most liked exhibit, visitors mentioned knowledge or learning in the second position for the exhibition but in a marginal position (after reasons related mainly to technology and communication) for VR systems. Moreover, in the specific questions about general usefulness, specific contribution and suitability for learning, visitors underlined VR's experiential character and capacity to provide a general visual image, while the exhibition, although being also very interactive ανδ experiential, was associated more with new, more detailed and well documented knowledge about historical facts, institutions and people.

The question about recall confirms that different exhibits provide different kinds of learning (more spatial, visual and superficial in the VR applications vs more factual, abstract and specific in the exhibition) and that people are more able to remember cultural contents in the Kivotos and the exhibition than in the Tholos. This would justify the apparent "conflict" between the known traditional way of learning (associated to text and exhibitions) and the experiential character of new media, which obviously also transmits knowledge but probably does so in a different way than the one that visitors expect or are able to perceive.

It is possible that because of the specific conditions of the Hellenic Cosmos's exhibits (but it is extendable to all ICT applications in the CH field), interviewed visitors confounded different factors, and their answers with regard to the real suitability of ICT for CH settings are blurred. Visitors considered that the most suitable exhibit for learning was the exhibition because it combined different means of presentation (with an important presence of images and text), which could be explored freely in order to obtain factual or historical information. On the other hand, the technological exhibits, because of their audiovisual character, were considered suitable for obtaining spatial information. Yet, the difference in the effectiveness of the VR exhibits and the exhibition is more complicated than just the dichotomy visual/textual. It involves content (spatial vs historical), means of communication (audiovisual/experiential vs textual) and also control and interaction (almost non-existent at the VR exhibits). And, more importantly, these factors are different and independent of each other, although there is a link of suitability between the first two: textual means seem to be more suitable for conveying historical information, while visual means are more efficient at showing spatial and dynamic information (Economou and Pujol, 2006).

Therefore, as the case of the Kivotos demonstrates, the use of visual information does not necessarily imply that the learning related to it is more superficial. It is also a matter of content. The constructivist learning paradigm states that previous knowledge is paramount for learning because it determines the relevance of the new information, acts as a "previous organizer" for its caption and generates the "cognitive conflict" that allows conceptual change (Pozo, Asensio et al., 1989). This is what might have happened at the Kivotos, which despite being visual, displayed a known subject (the Olympic Games), which was also relevant (showed human activity and was connected to visitors' experiences) but presented differences with previous knowledge (between ancient and modern Olympic Games).

On the other hand, visitors' apparent incapacity to say anything specific they learned at the Tholos may also be due to a methodological issue related to the representation means: as previously said, factual content is more easily expressed verbally than spatial and procedural content, for which visual means seem more suitable. At Tholos, spatial elements are visually displayed but the only procedural information that takes advantage of VR's dynamic potential is the navigation around buildings and eventually the construction of the Attalus' Stoa. This might justify that visitors could not remember any specific spatial fact (a part from the impressive building events) or, at least, were not able to express these verbally because they had been stored in a visual way.

Finally, the last reason for our results may be the lack of direct interaction in the VR exhibits and especially in the Tholos: in their answers, visitors manifested that they had especially appreciated the possibility of participation and control at the exhibition and criticized the lack of these features at the Tholos. Some authors (Dale, 1969) have demonstrated that we remember better what we see than what we hear and even better what we do. Therefore, it is possible that, although historical information was provided at the Tholos, its poor results in comparison with the exhibition are also due to the passive role adopted by visitors, since all the interaction with the system was carried out by the museum educators and the former only participated at the final voting (positively mentioned by 9.2% of them). Moreover, because this information was orally superimposed to the navigation and had to compete with the visual input, which in our cognitive system has always a stronger impact, it might have been superficially stored and more quickly forgotten.

In conclusion, and in spite of its novelty and impressive technological investment with regard to visual realism and immersivity, the Tholos seemed to obtain the poorest results with regard to its main goal -learning- for four reasons: because of the low relevance of the content in comparison with the exhibition and especially with the Kivotos (lack of human presence in the application, unknown subject); because it only presented spatial navigation (visual information stored and therefore not retrievable by verbal means); because of the lack of direct interaction of visitors with the system (in comparison with the Kivotos and especially with the exhibition); and finally because historical information was presented orally (and might have consequently been stored marginally in comparison with the visual input). In this sense, and given the strong impact of Tholos' technological dimension on visitors (expressed in several answers to the questionnaire), this exhibit constitutes the most evident example that VR still suffers from "media hyperbole" (Roussou, 2000) and therefore fails to achieve its full potential as an effective communication mediator in CH settings (where it is supposed to convey knowledge through a transparent, intuitive interface).

On the other hand, this leads us to discuss one of the main benefits of ICT invoked by visitors and guides: that because they are less abstract than books they are considered good for CH settings and for children. Taking into account the previous argumentation, we could not be sure that visual means are *per se* the best option for children, especially without interaction, because real learning -understood as the long-term acquisition of conceptual, procedural and attitudinal knowledge (Pozo, Asensio et al., 1989)- can only be achieved through cognitive "effort", that is, a deep and conscious engagement with the content and the process of learning. Moreover, in the case of children, learning is not only about pure memorization of factual content but mainly about developing social and reasoning skills that will configure their competence as adults. In this respect, some cognitive studies (Calvert, 2002; Knipp, 2003), have demonstrated that technological applications are not very suitable because learning can only be achieved through systematic training involving abstract means, social interaction, mental representations (imagination) and the whole body, while electronic games and computer-learning environments focus almost exclusively on visual aspects, reduce the need for mental representations and tend to isolate the user.

With regard to CH settings, since these are informal learning environments, their learning goals are traditionally more related with contents and enjoyment than with the construction of reasoning skills. However, considering the effectiveness of texts and hands-on exhibits evidenced by this and other surveys, as well as the problems of integration of ICT applications within the exhibition design (Jovet, 2003) and the social dimension (vom Lehn, Heath et al., 2005; Pujol and Economou, 2006) provoked by ICT's computational interface and language, can we be sure of their suitability beyond their demonstrated positive effects over attention, motivation and engagement in both formal and informal environments (Pimentel and Teixeira, 1995; Wheeler, Waite et al., 2002)?

We are not yet able to answer categorically this question. We would be inclined to say, given the current amount of empirical knowledge, that the usefulness of ICT is not a matter of age or context but of content and communication: because of their virtuality and audiovisual character, they are useful to combine realistic and schematic elements aimed at showing spatial elements, discovering patterns, emphasizing relevant aspects of dynamic phenomena and making abstract concepts more concrete (as demonstrated by the Agora exhibition's big touch screens). Their reconstructive, immersive and interactive capacity can achieve a sensation of presence that would be suitable to learn about intangible aspects of cultures through game and discovery strategies (Pujol and Champion, 2007). On the other hand, when they are interactive, they overcome the traditional uni-directionality of mass media and allow visitors to modify the content in order to learn through an experimental process or to express their opinion/interpretation of the subject (as demonstrated by the recording screen at the Agora exhibition), which would be coherent with the social function of heritage and its settings.

4. CONCLUSIONS

The work presented in this paper is part of a larger project aimed at evaluating the suitability of ICT for CH with regard to learning, exhibition design and social interaction, as well as their perception and use by audiences. Here we presented an investigation of the effectiveness of immersive VR for learning about archaeology and the past.

The first conclusion of the survey is the general perception by visitors that all technological exhibits at the Hellenic Cosmos offer a similar feeling of transportation, participation or experience of the past, which is aimed at learning and which makes them better in their minds than books. Taking into account the results of previous visitor surveys who investigated that question (Owen, Buhalis et al., 2005; Forte, Pescarin et al., 2006), we would conclude that this is representative of the general perception of ICT applications in CH by audiences. It is also worth noting that at the Hellenic Cosmos we found a significant association with gender: women seemed to look more for learning than men, who were more interested in the technological dimension.

From visitors' opinions and recalls, it appears that each exhibit support a different kind of learning. The Tholos and the Kivotos were associated with more experiential outputs (which in the case of the Kivotos became a real feeling of transportation) and considered suitable for obtaining a global idea of spatial details. This was due to their audiovisual tri-dimensionality, immersivity and the resulting effects of conscious and unconscious presence. On the other hand, the Agora exhibition was considered better for learning because it contained factual/historical knowledge, was richer in communication means (with text being the basic one) and allowed interaction and control over the visit.

This evidences a contradiction between the declared purpose of ICT in the CH field (learning) and the perceived / objective outcomes (experiential aspects, difficulty to recall specific and especially historical contents). The perceived outcomes can be justified by the experiential character of new media, which obviously convey knowledge but probably do so in a different way than the one that visitors expect, which is associated to text, exhibitions and factual knowledge. The recorded outcomes are probably due to the fact that ICT's effectiveness is not a function of technology alone but a combination of several factors related with the content (spatial vs factual information, relevance of the subject); the communication strategy (visual versus textual means, coherent methodology of recall); and interaction (active exploration of the resources). This allows us to provide some guidelines for future applications: in order to be suitable for learning, CH virtual worlds need to be complete and show not only a visually realistic reconstruction of architecture but a real interactive and meaningful reconstruction of the past, containing active human presence.

And this leads to the final conclusion of the survey, which was related with the common believe that VR is more suitable for children and CH because its experiential and visual character makes it "easier" and more enjoyable. Our results confirmed those of previous studies concerning learning processes and integration within exhibitions, showing that the usefulness of ICT should not be related with age or context (because from these perspectives, it presents more disadvantages than benefits) but with content and communication. VR's experiential character (if there is full intuitive physical and social interaction) makes it suitable for learning about CH through the feeling of transportation into another culture. However, this has to be further investigated because the introduction of Presence in the CH field is very recent. Other applications (aimed at conveying historical knowledge) need to be supported by verbal explanations and direct interaction in order to effectively remain in users' mind. On the other hand, VR's visual and dynamic character makes it suitable for spatial phenomena; and its interactive potential makes it suitable for discovery learning and bi-directional communication.

Therefore, the answer to the title's question would be that, with regard to learning in CH settings, VR is not always worth a thousand words; only when the aforementioned specific communication and thematic advantages are useful for the exhibition's purposes and compensate the current drawbacks imposed by its computational interface and language. In this sense, we need to continue investigating systematically the effectiveness of old and new applications, and on the other hand, to experiment with new interfaces, designed from a museological perspective instead of a computational one. The VR multi-user application at the Hellenic Cosmos exhibition and a previous study (NG, 2002) demonstrate that this is possible from the point of view of the physical shape of the exhibit, which can be adapted to support social and museographical integration; but the communication language still poses a major challenge with regard to intuitiveness of use. And this is critical because the software constitutes the real link between the content and the user (Gaitatzes, Christopoulos et al., 2000).

The comparison of the different kinds of communication solutions at the Hellenic Cosmos's exhibition might provide some very general guidelines for the "perfect" exhibit (suitable for learning, enjoyable, engaging, easy to use, and suitable for groups), which have remained mostly the same since the first were proposed intuitively by the Foundation of the Hellenic World's researchers (Roussou, 2000). Firstly, it needs to contain different linked elements, including text (for learning) and curious / relevant things (for engagement). Secondly, it needs to be interactive (manipulation makes it enjoyable and suitable for learning) and intuitive (user-friendly interfaces with clear affordances, in order to concentrate on the content). Finally, it needs to allow exploration in groups (to be engaging) and integration within the exhibition context.

Technological exhibits are not yet able to provide all these elements together, especially with regard to exploration in groups and intuitiveness. This is due, as previously said, to the limitations of the interfaces, and especially of the communication language. At the same time, interaction is also very restricted and consists only of spatial navigation, even though the essential definition of VR makes it suitable for two possible educational applications: real participation in historical events (Presence) or manipulation of archaeological data (discovery learning). We believe that tangible or mixed reality interfaces as well as serious games technology are two of the most promising trends worth exploring more in the future because they allow all the features mentioned here with regard to interaction, group exploration and usability.

5. REFERENCES

Belaën, F., 2003. L'immersion Au Service Des Musées De Sciences. IN: *Ichim03: Cultural Institutions and Digital Technology*. Paris, Ecole du Louvre.

Calvert, S. L., 2002. The Social Impact of Virtual Reality. http://vehand.eugr.ucf.edu/handbook/Chapters/chapter38.pdf (accessed 27 Aug. 2004).

Dale, E., 1969. *Audiovisual Methods in Teaching*. New York, Holt, Rinehart & Winston.

Economou, M. and Pujol, L., 2006. Educational Tool or Expensive Toy? Evaluating VR Evaluation and Its Relevance for Virtual Heritage. In: *New Heritage: Beyond Verosimilitude. Conference on Cultural Heritage and New Media. Proceedings of the New Heritage Conference, Hong-Kong, 13th-14th March 2006*. Kvan, T. and Kalay, Y. Hong Kong, The Faculty of Architecture, University of Hong-Kong, pp. 284-302.

Forte, M., Pescarin, S., et al., 2006. VR Applications, New Devices and Museums: Visitors' Feedback and Learning. A Preliminary Report. In: *The E-Volution of Information Technology in Cultural Heritage. Where Hi-Tech Touches the Past: Risks and Challenges for the 21st Century. VAST06.* Eurographics, Nicosia, Cyprus, Short papers volume, pp. 64-69.

Gaitatzes, A., Christopoulos, D., et al., 2000. Hellenic Cultural Heritage through Immersive Virtual Archaeology. In: *Proceedings of the 6th International Conference on Virtual Systems and Multimedia, October 3-6, 2000*. Ogaki, Japan, pp. 57-64.

Jovet, V., 2003. Le Multimedia Dans L'exposition: La Double Problématique De L'appropriation Et De L'intégration D'un Media Marginal. In: *Ichim03: Cultural Institutions and Digital Technology*. Paris, Ecole du Louvre.

Knipp, T., 2003. Design Education in the Era of Technology: Considering Visual Perception. In: *Ichim03: Cultural Institutions and Digital Technology*. Paris, Ecole du Louvre.

NG, K. H., 2002. Mixed Reality Interaction for Group Experience. M. Sc. Dissertation, School of Computer of Computer Science and Information Technologies, The University of Nottingham. Nottingham.

Osberg, K. M. (1997). Constructivism in Practice: The Case for Meaning-Making in the Virtual World. Department of Education, University of Washington, Washington.

Owen, R., Buhalis, D., et al., 2005. Visitor's Evaluations of Icts Used in Cultural Heritage. In: *Vast 2005: 6th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*. Mudge, M., Ryan, N. and Scopigno, R. Europgrahics, Pisa, pp. 129-136.

Pimentel, K. and Teixeira, K., 1995. *Virtual Reality: Through the New Looking Glass*. New York, McGraw-Hill.

Pozo, J. I., Asensio, M., et al., 1989. Modelos De Aprendizaje - Enseñanza De La Historia. In: *La Enseñanza De Las Ciencias Sociales*. Carretero, M., Pozo, J. I. and Asensio, M. Visor, Madrid, pp. 211-240.

Pujol, L. and Champion, E., 2007. A Critical Examination of Presence Applied to Cultural Heritage. In: *Proceedings of the 10th Annual International Workshop on Presence*. Moreno, L. Starlab, Barcelona, pp. 245-256.

Pujol, L. and Economou, M., 2006. Evaluating the Social Context of Ict Applications in Museum Exhibitions. In: *The E-Volution of Information Technology in Cultural Heritage. Where Hi-Tech Touches the Past: Risks and Challenges for the 21st Century. VAST06*. Eurographics, Nicosia, Cyprus, Full Papers Volume, pp. 219-228.

Pujol, L. and Economou, M., 2007. Exploring the Suitability of Virtual Reality Interactivity for Exhibitions through an Integrated Evaluation: The Case of the Ename Museum. *Online International Museology Journal*, 4, pp. 84-97.

Reid, D. J., Zhang, J., et al., 2003. Supporting Scientific Discovery Learning in a Simulation Environment. *Journal of Computer Assisted Learning*, 19(1), pp. 9-20.

Roussou, M., 2000. Immersive Interactive Virtual Reality and Informal Education. In: *Proceedings of I3 Spring Days 2000 - Workshop on Interactive Learning Environments for Children.*

Schuemie, M. J., Van der Straaten, P., et al., 2001. Research on Vr: A Survey. *Cyberpsychology and Behavior*, 4(2), pp. 183-201.

vom Lehn, D., Heath, C., et al., 2005. Rethinking Interactivity: Design for Participation in Museums and Galleries. I: *Re-Thinking Technolgy in Museums: Towards a New Understanding of People's Experience in Museums*. Ciolfi, L., Cooke, M., Bannon, L. and Hall, T. Limerick, Ireland.

Wheeler, S., Waite, S. J., et al., 2002. Promoting Creative Thinking through the Use of Ict. *Journal of Computer Assisted Learning*, 18, pp. 367-378.

6. ACKNOWLEDGEMENTS

The research presented in this paper is part of CHIRON: Cultural Heritage Informatics Research Oriented Network, 2005-2008, which is supported by the European Community's Sixth Framework Programme under contract number MEST-CT-2004-514539. It has also been supported by the Greek General Secretariat of Research and Technology.

We are grateful to several staff members of the Foundation of the Hellenic World: Dimitris Efraimoglou (Managing Director), Giannis Paschalidis (Exhibition Planning Coordinator), Alexandra Xintara (Head of Museum Educators Dept.) and all the museum educators. Our thanks are also due to Panagis Koutsokostas, Antonia Chouvarda and Elvira Bogiatzi, MSc students in Cultural Informatics at the University of the Aegean in 2007, who carried out part of the data gathering at the Hellenic Cosmos.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

VR-BML: BEHAVIOUR MODELLING LANGUAGE FOR CONFIGURABLE VR APPLICATIONS

Krzysztof Walczak

Department of Information Technology, Poznan University of Economics - walczak@kti.ae.poznan.pl

KEY WORDS: Virtual Reality, Behaviour Modelling, X3D, VR-BML, Flex-VR, Cultural Heritage

ABSTRACT:

Creation of complex behaviour-rich and meaningful content is one of the main difficulties that currently limit wide use of virtual reality technologies in everyday applications. To enable widespread use of VR applications new methods of content creation must be developed. In this paper, we propose a novel approach to designing behaviour-rich virtual reality applications, called Flex-VR. The approach enables building configurable VR applications, in which content can be easily created and modified by domain experts or even common users without knowledge about VR design and computer programming. The VR content is configured from reusable programmable content elements, called VR-Beans. Appearance and behaviour of the VR-Beans are controlled by scripts programmed in a novel high-level language, called VR-BML (Behaviour Modelling Language). The language enables specification of generic behaviours of objects that can be dynamically composed into virtual scenes. The paper introduces the Flex-VR component and content models, describes the VR-BML language and provides an example of a Flex-VR application in the cultural heritage domain.

1. INTRODUCTION

Virtual reality technologies have reached the level of maturity that makes possible their use in a diversity of real-life applications. Cultural heritage, education, training and entertainment are notable examples of application domains that can largely benefit from the use of VR technologies – both immersive VR and desktop-VR.

Wider use of VR applications has been recently enabled by several factors: significant progress in hardware performance including cheap 3D accelerators present in almost every contemporary graphics card, development of platform-independent 3D content description standards, and availability of broadband multi-purpose communication networks that are able to deliver large amounts of data required by network-based VR applications. Increasing interest can be observed in exploitation of possibilities offered by lightweight desktop-VR applications implemented through 3D web content accessible on-line. Such applications do not require any special hardware or software investments and can be accessed both locally and remotely using a variety of devices.

However, despite the technical possibilities, the actual uptake of 3D applications on the web is still very low. One of the main problems that currently limit wide use of 3D applications on everyday basis is related to the creation of complex interactive behaviour-rich content.

Practical 3D web applications require enormous amounts of complex interactive content. Moreover – in most cases – the content must be created by domain experts (e.g., museum curators), who cannot be expected to have experience in computer programming or 3D content design. However, only involvement of domain experts guarantees sufficient amount of high-quality usable content, which may contribute to wider adoption of 3D applications in everyday use.

This motivation lead to the development of the Flex-VR approach to building VR applications, in which content can be relatively easily created and modified by domain experts and common users (Walczak, 2008a, 2008b). In Flex-VR, the VR application content is dynamically configured from reusable programmable content elements, called VR-Beans. Appearance and behaviour of the VR-Beans are controlled by high-level behaviour scripts.

In this paper, we describe a novel behaviour programming language, called VR-BML (Behaviour Modelling Language), which is used to program the behaviour scripts. The language enables specification of generic behaviours of objects that can be later dynamically composed into technically correct virtual scenes without additional programming.

The remainder of this paper is organized as follows. Section 2 provides an overview of the state of the art in programming of VR content behaviour. Section 3 contains an overview of the Flex-VR component model. In Section 4, the Flex-VR content model is shortly presented. Section 5 provides a detailed description of the VR-BML language. Section 6 provides an example of a practical application of the Flex-VR approach and the VR-BML language. Finally, Section 7 concludes the paper and indicates future works.

2. STATE OF THE ART

Significant research effort has been invested in the development of methods, languages and tools for programming behaviour of virtual reality content. These approaches can be classified into three main groups.

The first group constitute scripting languages for describing behaviour of virtual scenes. An example of a scripting language designed for programming VR interaction scenarios is MPML-VR (Multimodal Presentation Markup Language for VR) (Okazaki et al., 2002). Another interesting example is

VEML (Virtual Environment Markup Language) based on the concept of atomic simulations (Boukerche et al., 2005). An extension to the VRML/X3D standards enabling definition of behaviour of objects, called BDL, has been described in (Burrows and England, 2005). Another approach, based on the concept of aspect oriented programming has been proposed in (Mesing and Hellmich, 2006). Ajax3D is a recently developed method of programming interactive web 3D applications, based on the combination of JavaScript and the Scene Authoring Interface (SAI) (Parisi, 2006).

The second group of solutions are integrated application design frameworks. Such frameworks usually include some complex languages and tools that extend existing standards to provide additional functionality, in particular, enabling specification of virtual scene behaviour. Interesting works include Contigra (Dachselt et al., 2002) and Behavior3D (Dachselt and Rukzio, 2003), which are based on distributed standardized components that can be assembled into 3D scenes during the design phase. However, this approach is based on the classical VRML/X3D dataflow paradigm and standard event processing making it difficult to specify more complex behaviours. Another solution, based on the use of distributed components accessible through web services has been proposed in (Zhang and Gračanin, 2007).

The common motivation for developing new scripting languages and content design frameworks, as those described above, is to simplify the process of designing complex 3D presentations. However, even most high-level scripts and content description languages become complex when they are used for preparing complicated 3D presentations. The third group of solutions try to alleviate this problem by using graphical applications for designing complex 3D content. Virtools is a good example of this type of tools (Dassault, 2008). Recent research works in this field include (Arjomandy and Smedley, 2004) and (Vitzthum, 2006). However, even if graphical specification of behaviours may be in some cases more intuitive than programming, users still must deal with complex diagrams illustrating how a scenario progresses and reacts to user interactions. Such diagrams are usually too difficult to be effectively used by non-programmers.

The approaches described above may be successfully used by 3D designers and programmers for preparing fixed 3D scenes. This, however, is not sufficient to enable widespread use of VR applications. Such applications require creation of large amounts of complex interactive content by domain experts, such as museum curators or teachers. Therefore, a method is needed for efficient creation of interactive 3D content going beyond the capabilities of content creation methods available today.

3. THE FLEX-VR COMPONENT MODEL

3.1 Overview

The Flex-VR approach uses a specific organisation of the VR content, called Beh-VR. In Beh-VR, a VR application is built of software components called VR-Beans. Technically, VR-Beans are objects, implemented as standard *Script* nodes conforming to a specific convention. Conformance to this convention enables combining arbitrary sets of VR-Beans into technically correct 3D scenes and provides means of inter-object discovery and communication. Beh-VR applications are fully compliant with existing 3D content description standards – VRML/X3D – and therefore can run in standard 3D browsers (X3D, 2004).

Each VR-Bean consists of at least one scenario script, an optional set of media objects, and an optional set of properties (Figure 1). The scenario script is the main element controlling each VR-Bean. Scenarios are programmed in a high-level XML-based programming language called VR-BML (Behaviour Modelling Language). Scenarios describe what happens when the object is initialized, what actions are performed by the object and what are the responses of the object to external stimuli. In some cases there may be several different scenarios controlling the VR-Bean depending on the presentation context.

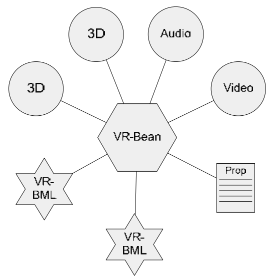


Figure 1. A VR-Bean object consisting of scenarios, media components and properties

A VR-Bean can have a number of media components, which are used for representing the VR-Bean in virtual scenes. Such media components may be 3D models (X3D/VRML), images, audio and video sequences, or texts. A VR-Bean can be also associated with a set of properties. Properties are characterized by a name, a type (integer, string, Boolean, etc.) and a value. Properties can be read by scenario scripts and can be used in determining appearance or behaviour of objects.

3.2 Scene Structure

A Beh-VR scene is created dynamically by combining independent VR-Bean objects. Each VR-Bean object is controlled by a VR-BML script (Figure 2). A behaviour script may load any number of media components into the virtual scene, thus creating geometrical, aural or behavioural manifestation of the VR-Bean – scene components. The scene components may be created during the object initialization phase or later during the object lifetime. Objects may also freely change their representations at any time. A scenario can create and destroy scene components and can communicate with the scene components within a single VR-Bean by sending and receiving events to/from the components. Each object can control all scene components created by the object, but has no direct influence on other scene components.

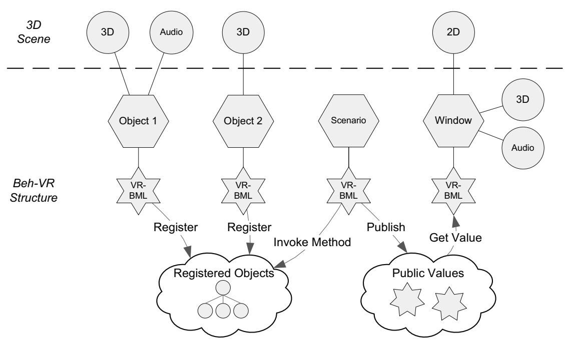


Figure 2. A Beh-VR structure and the resulting 3D scene

Since the contents of a Beh-VR scene is generally not known at the design time, communication between objects becomes a critical element. Meaningful communication requires identification of objects present in the scene, well-defined roles of objects and existence of technical means of communication.
Identification of objects is possible due to a process of registration and discovery. Each object may be registered in an arbitrary number of hierarchically organized categories. The categories also define roles of objects in the scene. An object assigned to more than one category plays several different roles in a scene. For example an object may be a status indicator and a game activator at the same time.

Communication between objects is realized using two mechanisms: *public values* and *method invocation*. Public values are named public expressions that can use variables and events from a single VR-Bean. Each VR-bean can explicitly read public values, can be notified when such a value changes, and public values can be directly assigned to input events of scene components. Method invocation can be performed on single objects, lists of objects and the whole categories. A method consists of a list of VR-BML commands, which may change the state of a VR-Bean, alter its representation in the virtual scene, invoke other methods, etc. Each method may have any number of parameters. Formal specification of parameters is provided in the method declaration, while their actual values are set in a method call.

4. THE FLEX-VR CONTENT MODEL

A Flex-VR application consist of a hierarchy of *presentation spaces*. Each presentation space may contain a set of VR-Bean objects, one or more *content templates* and optional *behaviour templates*. The content templates are responsible for dynamic generation of virtual scenes, which in the simplest case consists in creating scene nodes corresponding to the VR-Beans. Behaviour templates may be used to dynamically create VR-BML scripts for generic behaviour of objects.

To simplify creation of complex VR presentations, each presentation space may be associated with a *content pattern*, which adds semantics and predefined roles to objects. A content pattern defines a tree of categories together with default scenarios for objects in the categories and a tree of subspaces. Objects assigned to a category inherit methods and actions from the pattern scenarios, but can also override the default implementation with some object-specific implementation. Categories defined in a content pattern act as software component interfaces. There are dependencies between the categories, but not between particular objects, which may be assigned to the categories.

Designing a Beh-VR application consist in creation of presentation spaces and assignment of content patterns to the spaces and then objects to specific categories in the patterns. A content designer can also set properties of the spaces and the objects (appearance, size, location, state, etc.). While a pattern defines the overall structure of the presentation, the objects assigned to categories by a content designer provide the actual implementation of particular methods and actions. The assigned objects also contain specific media components such as 3D models and sounds. Therefore, entirely different presentations can be created by assigning different sets of objects to presentation spaces with the same content pattern.

5. THE VR-BML LANGUAGE

5.1 Overview

The VR-BML language is based on XML, which is the de-facto standard for creating new languages in the domain of multimedia and particularly 3D content. Most new 3D content standards such as X3D, 3D XML, and COLLADA are based on XML. XML is easy to interpret, existing software may be used for parsing and processing the program code. Importantly, the structure of XML programs may be verified against their formal specification in form of XML Schema.

The VR-BML language consists of a list of commands, a specification of program structure in form of XML Schema and expression grammar for attribute values. There are two main types of commands: *structuring commands* and *instruction commands*. Structuring commands provide the necessary structure of a VR-BML script, which enables to correctly interpret the instruction commands. Examples of structuring commands are: *Scenario*, *Initialize*, *Method* and *Action*.

Instruction commands are VR-BML commands that perform some operations, such as loading an X3D model into the scene (*Load*), moving a component to a specific location (*Move*), and calling a method (*Call*). Most commands require parameters, which may be provided as constant values or expressions containing references to variables, public values, and events generated by scene components.

VR-BML uses a hybrid approach based on both declarative programming for high-level elements (e.g., event actions) and imperative programming for low-level elements (e.g., algorithm details). This hybrid approach enables VR-BML to take best of the two worlds enabling the programmer to concentrate on important elements, and leave the common elements to the Flex-VR framework.

5.2 VR-BML Scenarios

A scenario script contains three sections (Listing 1): the *initialization section*, which describes what happens when the object is loaded into the scene, a number of *action statements*, which describe what actions are performed by the object as a result of changes or events in the scene (time, scene properties, user interactions), and a number of *methods* that can be explicitly called by the same or other objects in the scene. Scenario scripts can inherit methods and actions from scenario classes that form inheritance hierarchies.

```
<Scenario extends=”…”>
 <Initialize>
   ...
 </Initialize>

 <Action name=”…” cond=”…” time=”…”>
   ...
 </Action>

 <Method name=”…" param="…" wait="…">
  ...
 </Method>

 ...

</Scenario>
```
Listing 1. VR-BML scenario structure

The main commands used to build VR-BML scenarios are described below.

5.2.1 Scenario (extends): The *Scenario* command is the main element of each VR-BML scenario script. Each *Scenario* command is executed separately in parallel to other scenarios (e.g., through the use of threads). The *Scenario* element can contain only tree types of elements: *Initialize*, *Action* and *Method* (see Listing 1).

5.2.2 ScenarioClass (name): The *ScenarioClass* command is used to create classes of scenarios. The structure of a scenario class is the same as the structure a scenario (see Listing 1). It may contain the initialization section, a number of action statements and a number of method definitions.

The initialization section defined in a class is executed before initialization sections of any scenarios or subclasses that inherit from this class. The methods and action statements are inherited by subclasses and scenarios, but can be overridden by method definitions and action definitions in subclasses and scenarios. Methods and actions are unambiguously identified by their name – there is no method overloading based on the parameter sets.

5.2.3 Initialize: The *Initialize* command forms the initialization section of a VR-Bean. The initialization section is executed once for each VR-Bean object when the object is created. The initialization section typically contains all elements that are needed to build the VR-Bean and to set-up its initial state. Typically, the section contains commands for loading the scene representation of a VR-Bean, registering the VR-Bean in the virtual scene, and setting initial presentation attributes. The *Initialize* command may contain all non-structuring commands of the VR-BML language (e.g., *Load*, *Set*, *Register* and *Activate*).

5.2.4 Method (name, params): Each scenario may contain an arbitrary number of *Method* statements. Each *Method* statement defines a method – i.e. a fragment of VR-BML code, which is executed when explicitly called by the same or another VR-Bean.

The name of the method is provided in the *name* attribute, while formal parameters of the method are provided in the *params* attribute. Each parameter has a default value, which is used when a method call does not specify value of this parameter. Example method definition is the following:

```
<Method name="move" params="dst=0,0,0">
   ...
</Method>
```

5.2.5 Action (name, cond, count, time): Each scenario may also contain an arbitrary number of *Action* statements. Each *Action* statement represents an action – i.e. a fragment of VR-BML code, which is executed asynchronously when a specific condition is satisfied.

Execution of an *Action* command starts when the condition defined in the *condition* parameter is satisfied. The maximum number of executions is defined in the *count* parameter. Count equal *-1* means that there is no limit. If the *condition* is still true, the action is triggered again after the time provided in the *time* parameter. If the *time* parameter is equal *-1*, the action is repeated immediately after the previous iteration. Example of use:

```
<Action cond="{$size lt 10}" time="1000">
   ...
</Action>
```

5.3 Content Patterns

Content patterns can be assigned to presentation spaces. Patterns define the overall structure of Flex-VR presentation spaces simplifying the design of particular presentations. A content pattern defines the following:

- a list of categories of objects that can be assigned to this presentation space together with their default scenario definitions;
- a list of subspaces of the presentation space together with their default content patterns.

The overall structure of a content pattern definition is presented in Listing 2, while the meaning of particular commands is explained below.

```
<Pattern name="…"/>

 <Category name="…"/>
  <Scenario inherits="…">
   <Method name="…">
    ...
   </Method>
   ...
  </Scenario>
 </Category>

 <Category name="…"/>
  <Scenario>
   ...
  </Scenario>
 </Category>

 <Space name="…" translation="…" … pattern="…"/>

</Pattern>
```

Listing 2. Content pattern structure

5.3.1 Pattern (name): The *Pattern* command defines the overall content pattern. It has just one parameter: *name*, which enables this pattern to be referenced by presentation spaces. The *Pattern* element may contain two types of elements: *Category* and *Space*.

5.3.2 Category (name): The *Category* command defines a category of presentation objects that can be assigned to a presentation space associated with this pattern.

5.3.3 Space (name, translation, rotation, scale, pattern): The *Space* command defines a subspace of the presentation space. The subspace is described by a *name*, transformation parameters: *translation*, *rotation* and *scale* relative to the current presentation space and default content *pattern*. The pattern can be referenced by name from a library of existing patterns or defined directly in the *Space* element.

In the following sections the VR-BML instruction commands, which can be used in the scenario and pattern definitions are described.

5.4 General Purpose Commands

5.4.1 Set (name, value): The *Set* command implements assignment of a value to a variable. The command has two parameters: *name* and *value*. Example of use:

```
<Set name="i" value="10"/>
```

5.4.2 Call (object, method, param, wait): The *Call* command may be used to call methods of the same VR-Bean or other VR-Beans, enabling synchronous communication between VR-Beans. The *object* parameter specifies the target VR-Bean. This may be a single VR-Bean, a list of VR-Beans or a category of VR-Beans. If the object parameter is omitted, the *this* value is used by default. The *method* parameter provides the name of the method to call. The *param* attribute provides the actual values of the method call parameters. VR-BML uses explicit naming of parameters in method calls, which results in more self-documenting code. Finally, the *wait* parameter, specifies whether execution of the current VR-BML thread should be suspended until the execution of the method finishes or not. If *wait* is true, the next VR-BML command will be executed when the method finishes, if *wait* is false, execution of the VR-BML commands continues while the method is being executed. Example of use:

```
<Call object="{$co[$i]}" method="resize"
param="size=2"/>
```

5.4.3 If-Then-Else (cond): The *If-Then-Else* statement is a typical statement found in many programming languages. The *If* element specifies a condition (a Boolean expression), which may be met or not. The element may contain only two other elements: *Then* and *Else*. If the condition is met, the contents of the *Then* element is executed, otherwise the contents of the *Else* element is executed. Example of use:

```
<If cond="{$co!=null}">
   <Then>
      ...
   </Then>
   <Else>
      ...
   </Else>
</If>
```

5.4.4 Print (value): The *Print* command enables to print a value on a system console (e.g., a web browser console). The *Print* command is mostly useful for testing and debugging purposes. Example of use:

```
<Print value="Position={$posValue}"/>
```

5.5 Loop Commands

5.5.1 Every (time): The *Every* command is a loop that may be used to repeat execution of a fragment of VR-BML code in given constant time intervals. The interval is provided in the *time* attribute in milliseconds. Example use:

```
<Every time="{1000*moveTimeInterval}">
   <Call ...
</Every>
```

5.5.2 ForEach (list, var): The *ForEach* command is a loop, which is repeated once for every object in a list provided as the command parameter *list*. The elements of the list are being taken sequentially and assigned to a variable with the name provided in the *var* parameter. The *ForEach* command is very useful to execute specific operations on lists of objects that are retrieved dynamically. For example, all objects with the creation date outside a given time scale should disappear from the virtual museum exhibition room. Example of use:

```
<ForEach list="{$objList}" var="currObject">
   <Call object="{$currObject} method="…" … />
</ForEach>
```

5.5.3 For (name, from, to, step): The *For* command implements a typical *for* loop. The loop control variable is defined by *name*, while the starting, ending and step values are defined by the *from*, *to*, and *step* parameters, respectively.

5.5.4 While (cond): The *While* command implements a typical *while* loop, repeated as long as a condition is satisfied.

5.6 Object Discovery Commands

Since the contents of a Beh-VR scene is composed dynamically after the behaviour scripts are programmed, there must be a mechanism of discovering what objects are in the scene. This is achieved using the process of registration and discovery.

5.6.1 Register (category, name): The *Register* command enables to register the current VR-Bean object in a category. Categories form a hierarchical catalogue, enabling VR-Beans to discover what other VR-Beans are in the scene and what are their roles.

The *Register* command has two parameters: *category* and *name*. The *category* parameter specifies a comma-separated list of categories, where the VR-Bean should be registered. The *name* parameter specifies the name by which the VR-Bean will be registered in the category.

```
<Register category="objects/cultural/artistic,
objects/animated" name="Mortar"/>
```

5.6.2 Retrieve (category, var): The *Retrieve* command is complementary to the *Register* command and enables retrieving the list of VR-Beans which are registered in specific categories. The *Retrieve* command has two parameters: *category* and *var*. The *category* parameters provides the list of categories. If a category is followed by a "+" sign, all sub-categories of this category are also taken into account. The *var* parameter provides a name of a variable, which will contain the list of retrieved VR-Beans. Example of use:

```
<Retrieve category="obj/cultural+" var="obj"/>
```

5.7 Communication Commands

5.7.1 Publish (comp, public, value): The *Publish* command enables to create a new public value based on events coming from a component. The *Publish* command is useful for creating components that somehow control the virtual scene. For example, a scale slider may be created and its scale value may be published directly using the *Publish* command. Whenever the setting of the slider changes, the public value is automatically changed. The public value is not created directly from the component's event but it re-calculated using expression specified in the *value* parameter. Names of events in the expression are denoted starting with the tilde (~) sign. The *public* parameter specifies the name of the public value. Example of use of the *Publish* command:

```
<Publish comp="slider" public="scale"
value="{(~translation_out+5)/10}"/>
```

5.7.2 Assign (comp, public, event): The *Assign* command may be used to directly assign a public value to an event controlling particular component of the VR-Bean. For example, a public value *scale_of_objects* can be assigned to *set_scale* event of components of all presentation objects. In this way, by changing the public value, the scale of all presentation objects in the scene can be changed.

The *Assign* command is similar to the ROUTE statement in VRML/X3D/MPEG-4 standards, with the main difference, that ROUTEs must be established from the source to all receiving objects, which are not known when the virtual scene is being dynamically composed.

```
<Assign comp="comp1" event="set_scale"
value="scale_of_objects"/>
```

5.7.3 Link (comp, event, target, method): The *Link* command enables attaching events coming from VR-Bean components to method calls in the same or other VR-Beans. Since the Beh-VR scenes are configured dynamically, the list of target VR-Bean objects may not be known. Therefore the *Link* command enables specifying either a list of concrete VR-Beans or a category. The *target* parameter may be a list of objects, a list of categories, or may be left empty (not specified). If *target* is empty, the current VR-Bean will be the target. The *Link* command has the following parameters:

- *comp* – is the name of the component, which is the source of events;
- *event* – is the name of the event activating the method;
- *target* – represents the list of target VR-Beans or the list of target VR-Bean categories. If not specified, the current VR-Bean is used;
- *method* – specifies the name and the parameters of the method being called.

Example of use of the *Link* command:

```
<Link comp="s1" event="touchTime"
method="moveTo(0,0,0)"/>
```

5.8 Component Management Commands

The following commands are used to manage components representing a VR-Bean in a virtual scene.

5.8.1 Load (uri, comp, active): The *Load* command enables loading components into the Beh-VR scene. A component may be a 3D object, an image, a sound or any other element which may be directly or indirectly manifested in the virtual scene. The *uri* parameter provides location of the component – this may be a remote object accessed by the *http://* protocol, a local file system component accessed by *file://* or an object retrieved from a database. The *comp* parameter specifies the name that will be assigned to the newly loaded component. The *active* parameter specifies, whether the component should be automatically activated after loading or should stay as inactive until it is activated with the *Activate* command. Example of use:

```
<Load url="{$path}/oid={$o}" comp="object_{$o}"
active="false"/>
```

5.8.2 Activate (comp, active): The *Activate* command activates a component that has been loaded (using the *Load* command) into the Beh-VR scene. Because the loading process may include content generation, transmission over the network and parsing, it may take relatively long time. Therefore, to allow Beh-VR applications to switch on and off particular components in an efficient way, the *Activate* command may be used. Example of use:

```
<Activate comp="wheel" active="true"/>
```

5.8.3 SetPosition (comp, value, relative): The *SetPosition* command is used to position a component in the 3D space. The command has three parameters: *comp*, *value* and *relative*. The *comp* parameter specifies the name of the component to be positioned. The *value* parameter provides the new position of the component, while the *relative* parameter indicates whether the new position is absolute or relative to the current position. Example of use:

```
<SetPosition comp="object_1" value="0,10,0"/>
```

5.8.4 SetScale (comp, value, relative): The *SetScale* command enables to set a new scale factor to a component. The command has three parameters: *comp*, *value* and *relative* with the meaning similar as in the *SetPosition* command.

5.8.5 SetOrientation (comp, axis, angle, relative): The *SetOrientation* command enables to rotate an object in the three-dimensional space. The *comp* parameter specifies the name of the component to be rotated. The *axis* and *angle* parameters specify the axis and angle by which the component should be rotated.

5.8.6 GetPosition (comp, var): The *GetPosition* command enables retrieval of the current component's position in the scene. The position is assigned to a variable whose name is provided in the *var* parameter.

5.8.7 GetScale (comp, var): The *GetScale* command enables retrieval of the current scale of the component.

5.8.8 GetOrientation (comp, var): The *GetOrientation* command enables retrieval of the current orientation of the component in space.

5.8.9 Send (comp, event, value): The *Send* command enables to send an event to a component. The *comp* parameter specifies the name of the component, to which the event should be sent. The *event* parameter specifies the name of the event, and the *value* parameter specifies the value of the event to be sent. Example of use of the *Send* command:

```
<Send comp="plate" event="set_color" value="0.2
0.2 0.8"/>
```

5.8.10 Read (comp, event, var): The *Read* command enables retrieving the current value of an output field of a component. The names of the component and the field are provided in the *comp* and *event* parameters, respectively. The *var* parameter specifies the name of the variable that will hold the retrieved value. Example of use:

```
<Read comp="door" event="isOpen" var="doorOp"/>
```

5.8.11 Unload (comp): The *Unload* command is complementary to the *Load* command and causes removal of the object from the Beh-VR scene (regardless whether the component is currently active or not).

5.9 Animation Commands

5.9.1 MoveTo (comp, position, relative, time, wait): The *MoveTo* command is one of the animation commands in VR-BML. Using this command it is possible to animate a component from its current location to a new position. The *comp* parameter specifies the component to be animated. The *position* and *relative* parameters provide the target location. The *time* parameter specifies how long (in milliseconds) the animation should last. Finally, the *wait* parameter, specifies whether execution of the current VR-BML thread should be suspended until the animation finishes or not. Example of use:

```
<Move comp=″o1″ position=″0,2,0″ time=″500″
wait=″false″/>
```

5.9.2 RotateTo (comp, axis, angle, relative, time, wait): The *RotateTo* command enables to animate rotation of a component. The *comp* parameter specifies the name of the component to be animated. The *axis* and *angle* parameters specify how the component should be rotated. The meaning of *time* and *wait* parameters is analogous to the corresponding parameters in the *MoveTo* command.

5.9.3 ScaleTo (comp, value, relative, time, wait): The *Scale* command is an animation command which enables to set a new scale factor to a component through an animation. The command has five parameters: *comp*, *value*, *relative*, *time*, and *wait* (see *MoveTo* command).

5.10 Sensor Commands

5.10.1 PlaneSensor (comp, orientation, active): The *PlaneSensor* command is one of four sensor commands available in VR-BML. The *PlaneSensor* command is an easy way to add interactivity to a virtual scene by enabling a component to be moved by a user on a two-dimensional plane. The name of the component is provided in the *comp* parameter. The *orientation* parameter provides the orientation of the two-dimensional plane in space. The *active* parameter specifies whether the plane sensor should be activated or de-activated. Example of use:

```
<PlaneSensor comp=″object_1″
orientation=″0,1,0,0″ active=″true″/>
```

5.10.2 SphereSensor (comp, active): The *SphereSensor* command enables to add interactivity to a virtual scene by enabling a user to freely rotate a component in space. The *comp* parameter specifies which component should be allowed to be rotated. The *active* parameter specifies whether the sphere sensor should be activated or de-activated. Example of use:

```
<SphereSensor comp=″object_1″ active=″true″/>
```

5.10.3 CylinderSensor (comp, orientation, active): The *CylinderSensor* command enables a user to rotate a component in space around an axis. The *comp* parameter specifies the component. The *orientation* parameter provides the orientation of the rotation axis in space. The *active* parameter specifies whether the sensor should be activated or de-activated. Example of use:

```
<CylinderSensor comp=″object_1″
orientation=″0,1,0,0″ active=″true″/>
```

5.10.4 TouchSensor (comp, method, active): The *TouchSensor* command enables to activate a component as a touch-sensitive element. The *comp* parameter provides the name of the component to become a touch sensor. The *method* parameter specifies the method invocation, which should be executed when the component is touched. The *active* parameter specifies, whether the touch sensor should be activated or de-activated. Example of use of the TouchSensor:

```
<TouchSensor comp=″object_1″
method=″playSound(id=1)″ active=″true″/>
```

6. APPLICATION OF FLEX-VR TO CULTURAL HERITAGE

The Flex-VR approach has been successfully applied in a virtual museum system called *ARCO – Augmented Representation of Cultural Objects* (ARCO). The system allows museum staff to setup virtual exhibitions from existing components by creating presentation spaces and by assigning content patterns, templates and content objects to the spaces (Walczak and Wiza, 2007).

The process of designing a virtual exhibition content in the ARCO system is presented in Figure 3. The content model consists of one presentation space (*Granary*) based on a content pattern (*Exhibition*). A content template is assigned to this space. The template contains a 3D environment model and a set of commands to create VR-Beans corresponding to content objects.

The content pattern defines categories of content objects and provides default implementation of category behaviour scripts. The pattern defines three categories: *Objects*, *Scenario* and *User Interface*. The categories are represented as folders in the application allowing a designer to easily assign objects to the categories.

Figure 3. Flex-VR content structure of an ARCO virtual museum exhibition

A presentation designer (e.g., a curator in a museum) can create different presentations based on the same content pattern by using different content templates and by assigning different sets of content objects to the categories, and setting their parameters and presentation properties (e.g., positions, sounds).

In Figure 4 an example view of the virtual exhibition is presented. A user navigating in a 3D virtual museum room can examine a collection of virtual museum objects. The user can interact with the objects and can activate them by pressing buttons. Activation of each object is programmed in the pattern, but behaviour of each object upon activation may be different as specified in the object's behaviour script. For example, the *Plane* object disassembles to show how it is constructed, the *Bells* object plays recorded bell sound, while the *Wooden Statue* object starts a voice description explaining particularities of the piece of art.

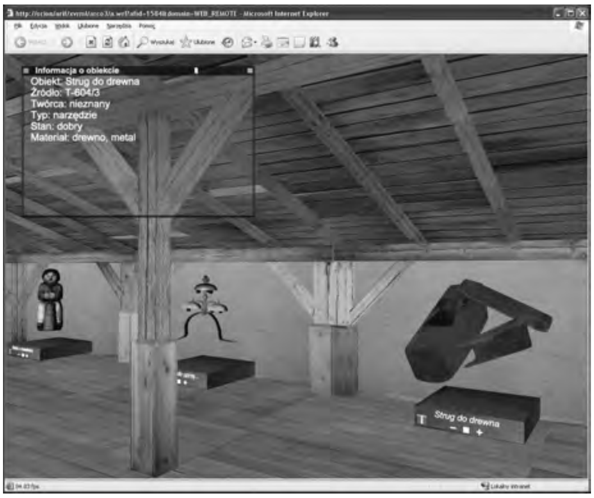


Figure 4. Flex-VR interactive virtual museum exhibition (Polish version)

7. CONCLUSIONS

The presented method of programming and composition of behaviour-rich 3D content simplifies the design of interactive VR applications in two ways. On the one hand it simplifies programming behaviour of virtual objects and scenes by enabling the use of the high-level VR-BML language instead of the low level elements such as sensors, interpolators, and routes. This language can be used by experienced users.

On the other hand, by providing the capability of dynamic content configuration, the Flex-VR approach enables users without programming experience to create complex interactive 3D content by combining predefined content patterns and collections of objects. Therefore, the Flex-VR approach can be used as a basis for building simple to use VR authoring applications for domain experts.

Future works include extending the VR-BML language with more advanced features, such as non-linear animation, complex objects and object groups. The authoring application will be also extended to support visual arrangement of objects.

REFERENCES

ARCO. Augmented Representation of Cultural Objects. http://www.arco-web.org/ (accessed 24 Aug. 2008)

Arjomandy, S., Smedley, T. J., 2004. Visual specification of behaviours in VRML worlds. In *Proc. of the 9th Int. Conf. on 3D Web Technology*, Monterey, CA, USA, 2004, pp.127-133

Boukerche, A., Zarrad A., Duarte, A., Araujo R., Andrade, L., 2005. A Novel Solution for the Development of Collaborative Virtual Environment Simulations in Large Scale. In *Proc. of the 9th IEEE Int. Symp. on Distributed Simulation and Real-Time Applications*, 2005, pp. 86-96

Burrows, T., England, D., 2005. YABLE-yet another behaviour language. In: *Proc. of the 10th Int. Conf. on 3D Web Technology*, Bangor, UK, 2005, pp. 65-73

Dachselt, R., Hinz, M., Meissner, K., 2002. Contigra: an XML-based architecture for component-oriented 3D applications. In: *Proc. 7th International Conference on 3D Web Technology*, Tempe, AZ, USA, 2002, pp.155-163

Dachselt, R., Rukzio, E., 2003. Behavior3D: an XML-based framework for 3D graphics behaviour. In: *Proc. of the 8th International Conference on 3D Web Technologies*, Saint Malo, France, 2003, pp.101-112

Dassault, 2008. Dassault Systèmes Virtools 4, http://www.virtools.com/solutions/products/ (accessed 24 Aug. 2008)

Mesing, B., Hellmich, C., 2006. Using aspect oriented methods to add behaviour to X3D documents. In: *Proc. of the 11th IC on 3D Web Technology*, Columbia, MD, USA, 2006, pp.97-107

Okazaki, N., Aya, S., Saeyor, S. And Ishizuka, M., 2002. A Multimodal Presentation Markup Language MPML-VR for a 3D Virtual Space. In: *Proc. of the Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenges*, Melbourne, Australia, 2002

Parisi, T., 2006. Ajax3D: The Open Platform for Rich 3D Web Applications, Whitepaper, Media Machines, Inc., Aug, 2006. http://www.ajax3d.org/whitepaper/ (accessed 24 Aug. 2008)

Vitzthum, A., 2006. SSIML/Components: a visual language for the abstract specification of 3D components. In: *Proc. of the 11th International Conference on 3D Web Technologies*, Columbia, MD, Apr 18-21, 2006, pp. 143-151

Walczak, K., 2008a. Dynamic Composition of Behavior-rich VR Content. In: *Proc. of the 1st International IEEE Conference on Information Technology*, Gdansk, May 18 – 21, 2008, ed. Andrzej Stepnowski, Marek Moszyński, Thaddeus Kochański, Jacek Dąbrowski, Gdansk University of Technology, pp. 333-336, ISBN: 978-1-4244-2244-9.

Walczak, K., 2008b. Structured Design of Interactive VR Applications. In: *Proc. of the 13th ACM Int. Symposium on 3D Web Technology – Web3D: Infrastructure Evolution 2008*, co-located with SIGGRAPH 2008, August 9-10, 2008, Los Angeles, CA, USA, pp. 105-113

Walczak, K. and Wiza, W., 2007. Designing Behaviour-rich Interactive Virtual Museum Exhibitions. In: *Proc. of the 8th Int. Symp. on Virtual Reality, Archaeology and Cultural Heritage VAST 2007*, Brighton (UK), Nov. 26-30, 2007, pp. 101-108.

X3D, 2004. International Standard ISO/IEC 19775:2004, *Extensible 3D (X3D)*, Web3D Consortium, http://www.web3d.org/ (accessed 24 Aug. 2008)

Zhang, X., Gračanin, D., 2007. From coarse-grained components to DVE applications: a service- and component-based framework. In: *Proc. of the 12th IC on 3D Web Technology*, Perugia, Italy, April 15 - 18, 2007, pp. 113-121

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

AN AUGMENTED 3D ALBUM BASED ON PHOTOS AND BUILDING MODELS

Pablo C. Elias [a], Asla M. Sá [a], Alberto Raposo [a], Paulo Cezar Carvalho[b], Marcelo Gattass [a]

[a] Catholic University of Rio de Janeiro, Computer Graphics Technology Group
R. M. de S. Vicente, 225, Rio de Janeiro, RJ, Brazil, 22453-900
(pelias, asla, abraposo, mgattass)@tecgraf.puc-rio.br
[b] IMPA, Instituto Nacional de Matemática Pura e Alicada, Rio de Janeiro - pcezar@impa.br

KEY WORDS: Calibration, Augmented Reality, 3D Photo Navigation, Image Processing, Camera Reconstruction

ABSTRACT:

Camera calibration is an important step in computer vision algorithms to interpret and reconstruct the three-dimensional structure of a real scene from a set of digital pictures or videos. Captured images are blended via computer vision techniques with synthetic images from scene models in order to render new applications, called augmented reality applications. Among the many uses for this technology this work has particular interest in those applications concerning augmented visits to buildings. These applications, produce images of buildings — typically old structures or ruins —reconstructed from virtual models inserted into captured pictures, allowing one to visualize the original appearance of those buildings. This work proposes an efficient, semi-automatic method to perform such reconstruction and to register virtual cameras from real pictures and models of buildings. This method makes it possible to compare photos and models through direct superimposition and to perform a three-dimensional navigation between the many registered pictures. Our method requires user interaction, but it is designed to be simple and productive.

1. INTRODUCTION

In photography collections on a same subject stored in data files, very often the file names carry little or no information about their specific contents. This makes storing and searching for pictures by content a difficult task.

With the possibility of 3D registration, many search functionalities for a given picture based on a region of the virtual scene can be developed in order to facilitate the management of image sets. The proposal of a three-dimensional picture filing system is not original, but what we propose herein is an application aimed at the cultural heritage context, focusing on the demands and the users involved in this field. For instance, in the case of incomplete buildings or ruins, it is possible to use a three-dimensional model of the finished building according to its historical registers, creating an augmented reality application in which real pictures can be compared with the complete virtual model.

The main functional requirements for the proposed application are: interactive positioning of virtual cameras based on a sparse set of pictures and computer models; ability to register the several reconstructed cameras in the synthetic scene; and finally, three-dimensional navigation among the registered cameras. The proposed technique assumes a known geometric model of the building and presents original techniques to position the pictures in relation to the model.

Camera reconstruction is one of the critical issues of computer vision. It consists in retrieving parameters that define the mathematical model of a camera (known as virtual camera) from the image of a real scene. Our method approaches the problem of computing the camera positions from a single picture, without using calibration markers applied to the building scene. To allow for a general position system the user must interactively navigate in the virtual model to provide an approximate camera position.

The proposed method uses integrated techniques that help the user create a set of corresponding features throughout digital images and the virtual model of a building. Based on these associations, the position of the camera that originated the real picture related to the model can be retrieved, thus providing navigation possibilities over the image set provided.

Although we propose a semi-automatic method, our basic rule is to demand little user interaction, minimizing the user's efforts. The interface concept proposed allows the user to handle *only the image* in order to perform the detection of associations with points of the world – direct associations between segments of the image and the model are not necessary. Nonetheless, camera reconstruction can also be made through a set of image-model associations explicitly provided by the user.

The test application implemented provides a retrieval solution and camera registration based on the semi-automatic matching between CAD models and pictures. The application provides tools to compare them, and a new three-dimensional navigation experience over the several pictures registered.

The case study presented involves a set of pictures and a model of a monastery from the 17th century, currently in ruins. However, as will be seen, there are no restrictions regarding the type of model used. Therefore, it is also possible to use models of objects whose structure is more precisely known, such as engineering or CAD models of modern buildings.

The paper is organized as follows: the next section summarizes the two main references related to this work and points out the differences in our proposal. In Section 3 the pre-processing steps required to the main method are presented. In Section 4 the camera position recovery technique is presented. Results obtained with our case study are shown in Section 5. Finally, conclusions and future work are presented in the last section.

2. RELATED WORK

Two photogrammetric systems have inspired the present work: Microsoft's Photosynth (Microsoft Live Labs, 2008) originated by Noah Snavely's work (Snavely et al., 2006), and Façade, proposed by Paul E. Debevec (Debevec, 1996).

Photosynth's input is a dense set of pictures of an object with overlapping regions. It retrieves clouds of three-dimensional points and camera models of these pictures based only on the matching features among them, which are calculated using the RANSAC (Fischer and Bolles, 1981) and SIFT (Lowe, 2004) algorithms. After matching features are located automatically, epipolar geometry is used to compute the fundamental matrix and retrieve world points, being refined according to the number of matches found. Even though this technique is successful, presenting efficient results especially regarding camera retrieval, it imposes an operational condition that is not desired in the present work: it assumes the availability of a dense set of pictures with overlapping regions, and it requires a long processing time.

Façade, on the other hand, adopts a semi-automatic approach in which a sparse set of pictures is used and the model's geometry is partially known. The camera model is reconstructed with the purpose of retrieving the proportions from a parametric model that was previously created by the user, associating several marks in several images. Debevec was based especially on (Taylor and Kriegman, 1995) to reconstruct the three-dimensional structure of a scene based on multiple pictures. He has also proposed a new method to estimate an initial camera rotation using orientations that are known in the scene. The camera adjustment technique and the method for matching models and images that we propose herein were motivated by such work.

In the initial stage of our research, the technique proposed by Debevec was considered. We concluded that it demands too much user interaction, conflicting with our requirements. Moreover, the final results of this technique are as good as the quality of the parametric model created and the marks made to the images. However, Debevec uses this technique with the principal goal of reconstructing the proportions of a rough model, which will subsequently be used as input for other integrated techniques, such as View-Dependent Texture Mapping (Debevec, 1996), depth map generation using the model as restriction, and other techniques that increase the quality of the final rendering – these are also outside our goals. The Façade system created by Debevec became the base for other commercial systems, and is successfully used to effectively retrieve model proportions and camera positions.

The fact that we assume the model's geometry to be known implies a substantial difference in relation to Photosynth and Façade's purposes, as well as allows a simplification of the camera reconstruction process. However, despite the different approaches, some concepts of Snavely's final application were used here to develop the test application, such as the calibration system using EXIF tags and the general concept of three-dimensional photo album.

3. PRE-PROCESSING

Two pre-processing steps are performed independently and concluded before the main method starts, namely: camera intrinsic parameter retrieval, and geometry loading and processing.

3.1 Camera Intrinsic Parameter Retrieval

A calibrated camera is one in which any given point **x** in its projection plane can be related to the respective ray **d** connecting its optical center to **x**. Formally, this means finding a calibration matrix **K** representing a transformation between this point **x** and the ray's direction, i.e., such that $\mathbf{d} = \mathbf{K}^{-1}\mathbf{x}$ (Hartley and Zisserman, 2003).

In this work, we have adopted the CCD pinhole camera model, which is similar to the classical pinhole model but taking into account that the pixels in the projection plane might not be square. In this case, the matrix **K** is given by:

$$K = \begin{bmatrix} \alpha_x & 0 & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

Entries in **K** are usually called the camera's intrinsic parameters. We refer to the reconstruction of those intrinsic parameters as the camera calibration stage. Such parameters are computed for each input image provided. If the image has proper EXIF information, the camera's geometry can be extracted directly from the image file metadata. If such information is not available, the calibration is made based on the vanishing points of the three principal directions.

Several authors (Rother, 2002; Schaffalitzky and Zisserman, 2000; McLean and Kotturi, 1995; Gamba et al., 1996) developed techniques for the automatic detection of vanishing points, but they only work in cases with little noise and in the presence of good straight line segments that can be detected in the image. Besides, the probability of false positives is high, i.e., often segments that do not belong to any of the three directions are detected, because straight lines that are perpendicular to the world can be projected as being almost parallel to the image.

Therefore, we have opted for a semi-automatic approach, with the principal directions of the scene being provided by the user, as illustrated in Figure 1. This picture was taken in the 1980's, so it hasn't EXIF tags and is very noisy, due to the digitalization process, representing a very difficult case for a full automatic approach.

Figure 1: Example of simple vanishing point calibration. Marked lines are indicated by the white arrows.

The method implemented for camera calibration is the one suggested by (Hartley and Zisserman, 2003). Although the vanishing point calibration provides a good estimation of the camera's orientation, only its intrinsic parameters are calculated at this stage. Camera orientation and translation will be obtained subsequently.

3.2 Structural Edge Extraction

Geometric meshes generated in modeling applications often present many edges that cannot correspond to any segment of the input image, either because they are not visible or because they result from mesh representation processes with rendering purposes rather than favoring an economical structural description of the model.

The virtual model pre-processing stage is comprised of a sequence of simple operations, resulting in a new list of edges called *structural edges* of the model. These are used to facilitate the location of matches between the model and the image. The pre-processing operations carried out on the original model are the following:

- Discarding duplicated edges;
- Discarding coplanar edges;
- Aggregating sequences of collinear edges;
- Normalizing the mesh and subsequently discarding short edges;
- Possibility of manually discarding edges.

Figure 2: Part of the original mesh of a model displayed in wireframe.

Tests of the simplification process using a model with over 1,000,000 edges resulted in a list with a little over 3,000 edges (0.3%), most of which will be discarded during the occlusion test that will be described below. In fact, in the model illustrated in Figures 2 and 3, a little over 90 visible edges were selected on average, to mark the relations between model and image.

Figure 3: Some of the selected visible structural edges marked in yellow on the model.

When a model's mesh is first loaded, it is processed and the list of structural edges is stored in an auxiliary file (.simp). The model's original geometry will be used again to enhance users visual experience and for the occlusion test stage described below.

4. CAMERA POSITION RECOVERY TECHNIQUE

To retrieve the point of view of a given camera, the identification of corresponding features between the picture and the virtual model is essential. The approximate virtual model of the scene captured in the images plays the role of calibration pattern for the cameras. This requires knowing its real proportions and processing the images so that the corresponding features between the images and the virtual model can be identified.

Since our case study is applied to buildings, the identification of edges rather than points was preferred, because they are easier to track in pictures of buildings. The approach adopted gives the user an important role manipulating the picture-model associations, especially where, due to reasons regarding the quality of the input image, it is not possible to detect a sufficient amount of good-quality edges.

The proposed method has the purpose of helping the user as much as possible, so that in simple cases where the image has little noise the user might not need to provide extra segments for the application to retrieve the camera. In these cases, the user will only need to provide one adequate initial camera position.

4.1 Initial Solution

The initial position of the virtual model, corresponding approximately to the point of view of the picture whose camera is to be retrieved, is assumed to be provided by the user. To position the model, the user enters manually, through a trackball-type manipulator, a translation and a rotation to be applied to the virtual camera. The goal is to align, as much as possible, the model's structural edges with the corresponding edges of the picture that is mapped onto the camera projection plane. From the observer's point of view, the image is always fixed in the screen space, while the model moves (Figure 4).

Figure 4: Model superimposed by an image with opacity of 100%. Even with maximum opacity, the guidelines are still visible and the model can be manipulated.

Using the user-provided initial position, the structural edges of the virtual model undergo an occlusion test that takes the original model into account, with the purpose of discarding structural edges which are not visible from the point of view of the initial position.

To implement the occlusion test, OpenGL's Occlusion Query functionality is used. The procedure is simple: the original model is rendered; then each of the structural edges is rendered; then the Occlusion Query verification is carried out. If the rendered edge generates visible pixels, then it is visible and must be maintained. Otherwise, the edge is discarded. The final result consists of a set V of visible structural edges.

4.2 Matching Structural Edges and Images

The semi-automatic method proposed to solve the problem of associating features of the picture to the virtual model consists of 3 stages: 1) edge detection on the image; 2) interpretation of the correspondence between the detected edges and the visible structural edges of the virtual model; 3) intervention and validation of matching features by the user.

4.2.1 Edge Detection on the Image

To locate straight line segments, Canny's filter is applied to the pictures to highlight the edges (Canny, 1986). Then, specific edge-detection algorithms are applied. We have evaluated the performance of two algorithms: the Standard Hough Transform (SHT) (Trucco and Verri, 1998) and the Progressive Probabilistic Hough Transform (PPHT) (Matas et al. 1998). The latter provides straight line segments on the image rather than straight line equations, as in SHT, as well as being significantly more efficient. The OpenCV library (Intel, 2008) was used for the implementation of both methods.

It was observed that each method presents a different performance depending on the input image quality, therefore it is not possible to single one out as the best in all cases. PPHT is more efficient when dealing with images with a significant amount of noise (such as the one in Figure 5), as was already mentioned; however, it is less precise. The choice for one of these two methods depends directly on the nature of the input image. The number of quality segments detected by PPHT was bigger than the test made with SHT, and the execution time was considerably smaller in high-resolution images.

Figure 5: Difficult case. Over 400 edges detected, most of them on the vegetation.

As a result of this stage, we obtained a set $U = \{u_{1,\dots,}u_n\}$ of edges from the image.

4.2.2 Matching Image and Structural Edges

The structural edges of set $V = \{v_1,\dots,v_n\}$ those that passed the occlusion test – are then projected onto the image space using the intrinsic camera parameters obtained in the pre-processing stage. Thus, a set $V' = \{v'_{1,\dots,}v'_n\}$ of straight line segments, given by the projection of the edges in set V onto the image space, is created.

Our idea is to use the model to restrict the area where the corresponding segments in the image will be searched, creating a window around each segment v'_i and thus defining a Region Of Interest $ROI(v'_i)$ relative to each edge. Each ROI works as a sub-image, restricting the search area for edges of set U that are candidate to match a structural edge v_i.

To select an edge from U that corresponds to an element within v_i, it is reasonable to analyze only those elements u_i contained in $ROI(v'_i)$, giving preference to elements with inclination, size and extremities comparable to those of V', because we assume the initial position of the model provided by the user to be approximately adjusted to the picture.

To objectively classify the edges of U that are matching candidates, i.e. that are inside the effective region of a ROI, we have adapted the error function proposed in (Taylor and Kriegman, 1995). Thus, each edge u_i from U inside $ROI(v'_i)$, is classified according to the following function:

$$rank(u_i) = \frac{l}{h_1^2 + h_1 h_2 + h_2^2} \tag{2}$$

where l is the size of segment u_i and h_1 and h_2 are the distances from the extremities of segment u_i to the straight line L_i, built from the endpoints of a projected model segment v_i (Figure 6).

The edge u_i with the best classification according to function *rank* is chosen as the one in the image that matches the structural edge v_i from the virtual model.

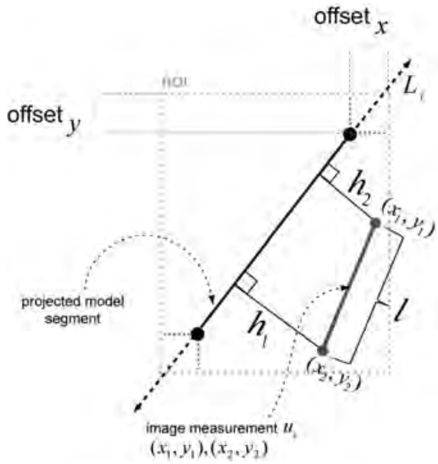


Figure 6: Edge classification within a ROI.

As can be observed, the *rank* function classification prioritizes large edges over smaller ones, but small edges that are closer and more aligned have priority over larger edges that are farther away or less aligned. The quality of the classification obtained depends on the position of the user-provided virtual model.

As this is a greedy local-search method, if there are intersections between *ROIs*, it might be the case that the same edge u_i in the image be classified as matching two or more different structural edges of the model, which is an inconsistent solution. To overcome this problem, the tie-breaker criterion adopted to select the structural edge that best corresponds to edge u_i in the image consists in comparing the value of *rank(u_i)* in relation to the different edges v_i classified as matching the same u_i, then again adopting a greedy strategy.

4.2.3 Validation and Manual Edition of Matching Features

The matching features between the model and the picture obtained automatically are finally displayed for the user using visual information to allow the user to discard them, approve them or complement them. The complementation entered by the user constitutes of new segments made only on the image. Based on these segments and on the initial camera position provided by the user, the best candidates to corresponding edges of the model are classified. Only in cases where it is impossible to detect matching features based on proximity (due to the low quality of the image), direct manual association between model and image is necessary.

Complementation is required when it is not possible to detect edges in the three principal directions of the scene, which is an essential requirement for the subsequent stage of camera adjustment (minimization). This can occur especially due to noise or occlusion. In this case, the user is required to complement the set of detected edges by marking new edges in regions where no or few segments were found.

In resume, the output of the matching step is a set of pairs *{(v_1,u_1), (v_2,u_2), ..., (v_n,u_n)}* to be used as input to the minimization step.

4.3 Optimizing Camera Position

Using the set of associations between model and image edges, already validated by the users, and the camera initial position, the next step is the camera position adjustment, in order to maximize the alignment between corresponding segments in the image and in the model.

To achieve this maximum alignment, it is necessary to measure the error between each pair *(v_i, u_i)*. With these measures, it is possible to recover the position and orientation of a calibrated camera, by means of the minimization of an objective function that depends on the external parameters and the obtained correspondences.

We use again the error function proposed by (Taylor and Kriegman, 1995), already presented in the previous section. That is related to the image formation process, modeled as a projection function **P**(**R**,**T**,v_i), where **R** is the rotation matrix and **T** is the translation matrix. The projection function produces an ideal bi-dimensional segment that is used to measure an error between the projected model and the user detected image lines u_i, given by:

$$O = \sum_{i=1}^{l} Error(\mathrm{P}(\mathbf{R}, \mathbf{T}, v_i), u_i) = \\ = \sum_{i=1}^{l} \mathbf{m}^T (\mathbf{A}^T \mathbf{B} \mathbf{A}) \mathbf{m} \quad (3)$$

where,

$$\mathbf{A} = \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix}, \mathbf{B} = \frac{l}{3(m_x^2 + m_y^2)} \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \quad (4)$$

and $\mathbf{m} = \{m_x, m_y, m_z\}^T$ is the normal vector shown in Figure 7.

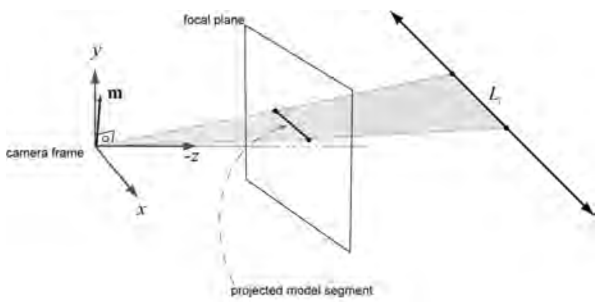


Figure 7: The representation of the line L_i by the normal vector **m**.

This function measures the total error between model and image edges in function of the camera external parameters, which are recovered using a minimization process to find their optimized values.

This objective function to be minimized is not linear. Therefore, in order to achieve the minimization, we approximate the non-linear function by a quadratic function and optimize it using a gradient-descent method.

5. RESULTS

5.1 Case Study

The set of pictures used as case study represents the ruins of the São Boaventura convent, which was built around 1660. The convent's model was created based on historic documents and photographic registers. The model is not totally precise in relation to the ruins' structure, but it suffices to illustrate and test the application and the method proposed.

The pictures employed include a lot of vegetation and noise, adding to the high complexity of this case study. Nonetheless, a set of 10 cameras was successfully retrieved from the respective pictures, in less than 10 minutes.'

Some results are shown in Figures 8 to 11:

Figure 8: A match between the model and a picture of the convent.

Figure 9: The same picture as in Figure 9, shown with zoom using a different orientation.

Figure 10: The model and its structural edges shown in yellow.

Figure 11: The match for the same camera position and orientation shown in Figure 11. The structural edges of the model are shown over the picture.

5.2 3D Photo Navigation

A test application was created to demonstrate the proposed method, performing semi-automatic retrieval of camera positions of pictures from a building in relation to its virtual model. Each retrieved camera can be registered, composing a picture filing and facilitating the management of the photographs in 3D space.

Moreover, the system provides a three-dimensional navigation experience over the pictures, offering a number of tools to compare details on the pictures and the model, such as zoom and transparency control.

The tridimensional navigation was implemented in a simple way to provide an efficient and easy mechanism for searching for pictures. Although such mechanism do not intend to be the best or the final solution for tridimensional navigation between camera positions, it can be used to illustrate the utility of the virtual camera recovering process. With such virtual cameras in hand, a new set of possibilities is opened allowing the creation of new features for picture filing, search and comparison between pictures and building models.

The tridimensional navigation algorithm receives as input a specific direction and the set of recovered cameras and produces and output that can be empty (in the case that no camera is sufficiently near) or represent another camera in the given set (chosen as being nearest in the given direction). Given a specific direction, one can navigate to another near camera in a specific direction using the criteria shown below:

$$dist(\hat{\mathbf{d}}, \hat{\mathbf{v}}) = \frac{1}{\langle \hat{\mathbf{d}}, \hat{\mathbf{v}} \rangle} \|\mathbf{d}\| \quad (5)$$

where *dist* is the distance classification function between a camera (i) in the given set and the current camera, $\hat{\mathbf{d}}$ is the unit vector that marks the direction to the optical center of the next camera i from the current camera and $\hat{\mathbf{v}}$ is the chosen navigation direction, represented in the local space of the current camera (Figure 12).

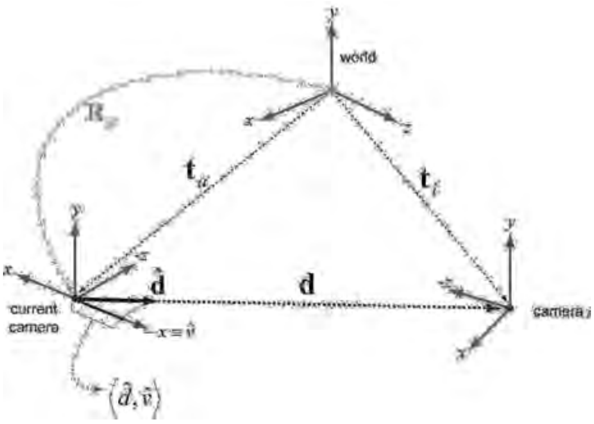


Figure 12: A simple strategy for tridimensional navigation between recovered cameras.

By using virtual cameras it is also possible to recover the position from where a picture was taken in the real world if one of the dimensions of the real model is known in one of the pictures of the used set. This allows one to build observations of a specific part of the model during the time (using the same world position). For example, an old picture can reveal that some parts of the real model are currently damaged or inexistent, as illustrated in Figures 13 and 14. If the position this old picture was taken is recovered, it is possible to take new pictures of the model from the same position, which makes possible to observe the same part of the model across the time.

Figure 13: A match between the model and a picture of the convent. The picture is shown with 100% of the opacity. This picture was taken in the early 80's so no EXIF info is available. Vanishing point calibration was used.

Figure 14: The same match shown in the Figure 13, using around 30% of opacity.

Another very useful functionality related to the production of archives for cultural heritage items is the possibility of verifying whether the set of pictures provided offers a complete view of a given building. To implement such functionality, one can simply project the pictures registered on the model by applying an inverse projective transformation using matrix $\mathbf{K}^{-1}$, and observe the parts of the model that are not covered by the pictures.

6. CONCLUSIONS AND FUTURE WORK

The present work has proposed a method for matching models and images of buildings using a set of integrated techniques, with camera reconstruction as the main strategy. To perform such reconstruction and successfully match and catalogue the pictures, first we needed to solve the problem of identifying correspondences between elements of the image and the model, which is one of the fundamental problems in computer vision. The approach proposed to solve it was to use the building's model, positioning it in order to restrict the search for matching features on the image. This strategy assumes the virtual model to be manipulated by the user in such a way that the edges can serve as guidelines to locate corresponding features in the image, using a local search strategy in the neighborhood of the projection of the model's edges.

The method is semi-automatic, beginning with an initial solution provided by the user which allows a local search for image-model associations rather than exploring the model's global information. If the input image contains significant noise and the photographed model has complex geometry, solving the matching problem becomes naturally difficult, and the method proposed herein becomes more dependent on user actions and prone to some degree of imprecision. In simple cases, on the other hand, the process is largely automatic and robust in relation to the model's initial position, as it is simpler to compute image-model correlations.

As final result, we have developed an application that implements the proposed method and provides a complete solution for the camera registration problem over pictures related to their virtual model. The system also provides various mechanisms to help the user compare pictures with the model, and navigate spatially over the several registered images.

Several improvements can be considered, especially in relation to the correspondence between points in the image and the world. In this sense, the use of NPR (Non-Photorealistic Rendering) techniques can be explored to detect structural edges in the virtual model and in the images.

A natural extension would be the application of this method to videos. The technique could be further developed to explore space and time coherence in frame sequences obtained from real videos. This extension to frame sequences can be facilitated by the very nature of the technique, which searches locally for corresponding features between model and image. When a sequence of frames from a film is assumed to present space and time coherence, the proposed method can be applied directly to adjust the sequence to the model, moving automatically from the initial frame position to the next frame. The process could begin with a manual positioning of the first video frame by the user, resulting in the retrieval of a complete camera path along a given time span.

As future work, we intend to use a set of pictures well distributed in space as input to a method that correlates them to a virtual model, seeking to improve the quality of an approximate virtual model of the represented building. The model's structural edges could be parameterized in order to be adjusted to the marks made to the pictures. This idea is very similar to the one developed by Paul Debevec (Debevec, 1996), but it would use models directly created by commercial modelers, which would benefit from the convenience of using a good modeler without demanding modeling precision. In other words, a model lacking the exact measures could be quickly created by a designer, and then its proportions would be adjusted based on the input pictures.

Another possibility to extend this work would be to use a dense set of pictures to reconstruct completely the correct proportions of a model, using only associations among the images. This strategy is inspired in the work developed by (Snavely et al., 2006) but it would use as base an imprecise model that could be adjusted and used as restriction for points reconstructed based on picture correlations using epipolar geometry. In (Snavely et al., 2006) only pictures are used, without any information on the model.

REFERENCES

Canny, J., 1986. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), pp. 679-698.

Debevec, P., 1996. Modeling and Rendering Architecture from Photographs. PhD thesis, University of California at Berkeley.

Fischer, M. A., and Bolles, R. C., 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM,* 24(6), pp. 381-395.

Gamba, P., Mecocci, A., and Salvatore, U., 1996. Vanishing Point Detection By A Voting Scheme. In: *IEEE International Conference on Image Processing*, Lausanne, Switzerland, pp. 301-304.

Hartley, R., and Zisserman, A., 2003. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridgeshire, UK, 2nd edition.

Intel, 2008. "Open Source Computer Vision Library". http://www.intel.com/technology/computing/opencv/ (accessed 16 Jun. 2008).

Lowe, D. G., 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision,* 60(2), pp. 91-110.

Matas, J., Galambos, C., and Kittler, J., 1998. Progressive Probabilistic Hough Transform. In: *British Machine Vision Conference*, pp. 256-265.

McLean, G., and Kotturi, D., 1995. Vanishing Point Detection by Line Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(11), pp. 1090-1095.

Microsoft Live Labs, 2008. "Photosynth". http://labs.live.com/photosynth/ (accessed 16 Jun. 2008).

Rother. C., 2002. A new approach to vanishing point detection in architectural environments. *Image and Vision Computing*, 20(9-10), pp. 647-655.

Schaffalitzky, F., and Zisserman, A., 2000. Planar Grouping for Automatic Detection of Vanishing Lines and Points. Technical Report, Department of Engineering Science, University of Oxford.

Snavely, N., Seitz, S. M., and Szeliski, R., 2006. Photo Tourism Exploring Photo Collections in 3D. Technical Report, Microsoft at Microsoft Research, Redmond WA, USA.

Taylor, C. J. and Kriegman, D. J., 1995. Structure and Motion from Line Segments in Multiple Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 17(11), pp. 1021-1032.

Trucco, E., and Verri, A., 1998. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, Inc, Upper Saddle River, NJ.

ACKNOWLEDGEMENTS

The TecGraf (Computer Graphics Technology Group) is one of the laboratories of the Computer Science Department at the Pontifical Catholic University of Rio de Janeiro (PUC-Rio) and is mainly supported by Petrobras. Alberto Raposo is also funded by CNPq, Process Number 472967/2007-0.

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

REAL TIME INTERACTION WITH CULTURAL HERITAGE OBJECTS IN VIRTUAL 3D

O. Domşa [a, *], M. Kadar [a]

[a] UAB, "1 Decembrie 1918" University, 510009 N. Iorga, 11-13, Alba Iulia, Alba, Romania
- (ovidiu, mkadar)@uab.ro

KEY WORDS: Virtual Heritage, Virtual Reality, Object interaction, Graphic engines, Open source, Programming, Landscape reconstruction, Real time

ABSTRACT:

Issues on Virtual Cultural Heritage using object interaction have been discussed in various circumstances and time spans, depending on hardware and software development. This paper proposes a new point of view regarding the development of software for Cultural Heritage and for learning purposes or knowledge framework proposed for children, young adults and large public, respectively. The application is structured in 3D reconstruction of landscape, graphic engines, programming interaction with the objects using voice or text, movement and interaction with haptic devices proposing a virtual historic educational game.
The modular structure of the project components allows modifying the content, the object interaction and the content of the dialog graph. We have studied the human virtual interaction in learning processes and the main components that offer interactions in the virtual learning environment. These mandatory guidelines are provided for preparation haptic interactions with objects in virtual environments. The virtual environments offer a good opportunity to change in real time some interaction factors like distances, speed and lighting. The application uses the free Irrlicht engine for rendering and sound interaction. The scenes contain different objects with different properties and different types of interaction like: text, sound or 3D moving objects. The application is organised into an XML database. By changing the landscape each object is reconstructed. The interaction with objects is defined by skyboxes with different forms.
The application provides different facilities including real time changing of lighting, camera position or distance between camera position and background.

1. INTRODUCTION

1.1 General 3D visualisation appreciation

The development off the graphic devices from 2D visualisation to 3D visualisation in the '80s has covered different areas of interest in sciences, simulation and Computer Games. New kids generation all over the world knows and tries different game systems. That means that 3D graphics and computer game simulation became a daily experience or part of the everyday life. Teachers in schools, visitors in museum, architects in reconstruction, engineering in simulation use 3D reconstruction and visualisation.
Different components of the environment pushed by a strong economic interest are strongly connected with entertainment, also to research and theoretical investments. Archaeology, History and Arts fields started their experiences in computer graphics and 3D visualisation in the same period (L. Calori, 2003).
We would like to underline the importance of the Virtual Reality systems even from the cultural point of view. The museums activities become closer than games, the kids and teens like to "play to learn", interaction and environments feeling become a necessity for attractive learning activities.
Historical sites from Romania are poorly presented using 3D methods, one of the main reasons being high performance in computer structures and graphic devices that means high costs in developing projects like that.
The projects high performance are well prepare, theoretical, to include also haptic devices in the project. We test the application with 3D glasses and haptic gloves.
We are also interest in simulations and special 3D archaeological graphic reconstructions.

1.2 Proposed aims

This paper proposes some theoretical and practical methods to create an interface between cultural heritage activities and game interactions. First of all we have created the landscape of the action, The Roman Citadel Apulum (Castru Roman Apulum, in Romanian language), the Roman 13-th Legion Fortress, placed on the Alba Iulia city II century A.C. with main buildings like they appear in the archaeology researches. Based on this reconstructed 3D environment we have placed some ancient people and artefacts created using 3D Max Studio or Maya applications. According with archaeological information about position or temporal placements of the ancient artefacts we have proposed a real time interaction using three different types of interaction: selection of the object and moving it using scale, rotation or translation, interaction with object using haptic devices and real time dialog using a graph dialog both with text and sound.

The application is created using open source software modules and free engines. The application was developed in Visual C# .NET. Virtual Heritage, in particular Archaeology Virtual Heritage of Roman Apulum Fortres, offers time-space connectivity with ancient history, interaction with objects like in virtual games, offering social and ancient life information, content feeling using haptic devices (for the fist experiments we used just haptic gloves).
The research aims is to offer an interactive, integrated, updateable, educational and open system for using in real time interaction with object in recreated virtual environments.

2. 3D MODELING, INTERACTION WITH OBJECTS

2.1 Historical consideration

We based our research on the Roman Citadel Apulum, placed in Romania, in one of the oldest cities in Transylvania, Alba Iulia.
The city is situated at an old gold and salt commercial crossroads, into the perimeter formed by the rivers Ampoi and Sebeº and the peaks of the Apuseni Mountains that mount mildly and lithely towards the terrace of the river Mureº and the Transylvanian Hills.
The gentle climate and the richness of the soil rendered this area habitable even since ancient times. Archaeologists register rich vestiges of the material culture - dating since Neolithic, Bronze Age, Hallstatt, Latene and Middle Ages - undeniable proof of our continuity on these territories.
The tribe of the Dacians from "the far-off Appulus" is mentioned in "Consolatio ad Liviam - Poetae latini minores", and the geographer Ptolemaios revealed in his "Geographical Guide" (written in the first half of the second century) the coordinates of the city: 49°15' longitude - 46° 41' latitude.
The XIII Geminae Legion is to be billeted here in one of the major Roman stone camp during the years Dacia was a Roman province. Along with the Dacians, the new comers (the Romans), "ex toto orbe romano", are the ancestors of the Romanian people, appropriating the Dacian ancient toponym Apoulon (a fortress situated at Piatra Craivii, 20 km North of Alba Iulia, which became the Roman Apulum).
Two roman cities, first Municipia and later Colonia, have developed near the Roman camps, into the fortress, but also nearby the Mureº river, in Partoº.
The settlings became two of the most wealthy and important places of Dacia - ("Chrysopolis" 251-253 d. Chr.) - outstanding in diversity and the novelty of the local civilization.
The case study is based on the archaeological research in the south east half part of the Mureº river terrace, occupied in the II-th Century BC by the Citadel and Canabae's of the XIII Roman Legion Geminae. The Apulum Citadel has benn constructed in 107-108 BC having military reasons but also strategic, economic and commercial aims. The first stage of the Citadel was only earth-like constructions having on the base 8,7 m width and 3,7 m high, after that stone Citadel was constructed. The calcareous stones that were used lately were brought here from ª ard, Bãrãbanþand Ighiu stone quarries.
They started the construction with the South Gate, *Principalis Dextra*, 10,2 m opening, double entrance, each of 4,2 m separated by the 1,8 m wall. The entrance is two towers flanked with 8,4 m x 6,6 m inside and 8,2 m x 6,3 m outside (Moga V., 1999).

Recent discovered new sites prove that it had rectangular form, 440 m on the east-west side and 430 m its north – south side with the similar orientation like common roman citadels from Inchuthil (piñata castra, 477/463 m.) in Bretany, Bonn (Bonna, 520/515 m.) in Germany or Vindonissa (Windich) in Switzerland and Vindobona (Wien) in Austria where XIII Geminae Legion pass trough in the first century (Moga V., 2002, p. 45).
The *Praetoria* and *Decumana* Gates were placed in middle of the sides, *Principalis Dextera* and *Sinistra* Gates were 2/3 of south respectively north sides. In Figure 1 we have presented a part of "Pianta D'Alba Iulia" the first and the oldest map of the city designed in the XVII century by an unknown Italian engineer, in witch he placed two important buildings of the Citadel in the right up site inside the castle and also the buildings placed around, A - Castello, B - Grand Chiefa, e Palaszo del Principe, C – Citta (down town) and D – Chiefa de Greci at the northern part of the Citadel.

Figure 1. A section of the "Pianta D Alba Iulia", XVII century (Kovacs A., 1984, Anghel Gh., 1985)

Archaeological research on the Apulum site was difficult because of successive overlap in building constructions, on the same site at least three different citadels were built in the Middle Age and feudal period. Starting with 1889, when Adalbert Cserni made the first digging in the area he find the big blocks of stone, 1,5-2 m long and 0,80 - 1,5 m high in the nearest of the south gate, Porta Principalis Sinistra (Moga V., 2002, p. 47).
The important researches in the area were started in 1966-1968 when important utilitarian constructions were started in the area. The other three gates were discovered and also observation towers, the rectangular houses ("bordei" in Romanian), buildings walls were constructed in *opus quadratum* style and *emplecton* (the core) in *open signinum* technique. The wall height was preserved is 6-7 m and three abutments were found in the exterior of the wall. The wall breadth was 2,1 m. Some stones were Aurelius Godes worker inscription, huts, ceramics, epigraphic monuments, coins chronologic evaluated from Hadrian to Gordian III and different sculptural monuments.
Building materials were carefully analysed starting with land construction, stone dimension and construction style for texture composer. Also we develop a long research in museum exposures to take photos and prepare good pictures for the original bricks, tiles, pavements or stone pipes.

Temples and polychrome mosaics, thermal baths, statues, amphitheatres, porticos, the governor's palace "Daciarum Trium" - that would be in brief the synthesis of this important military-political, economic-commercial and cultural-artistic centre, the miniature copy of the mother Rome.

2.2 Modeling and 3D reconstruction

The historical materials have been carefully analysed and has started the reconstruction with students from Computer Science speciality together with historical researcher working in the project "3D Virtual Museum", financed by the European Union, through "The National Cultural Found Administration", in 2007 - 2008.

3D objects modelling is well known that may be a determined step in landscape reconstruction.

The modelling has to decide between low-poly and high-poly object reconstruction. Our purpose is to offer real time interaction in this environment that we choose to represent the objects in low-poly format. Because our interest it was in real time interaction with the objects and our platform use different landscapes or environment reconstructed we present the archaeological studies in short, mentions that all the reconstruction was made by the rigorous documentation in collaboration with National "Unirii" Museum from Alba Iulia coordinated by Moga V. and Lascu I.

We chose to reconstruct the Roman Citadel, Roman Temple, Roman huts, cereal hangars, market places, roads, roman houses and head quarter villa. In the building reconstruction we keep the balance between efficiency, visual aspect, UV-layout, dimensions and scalability. Each object in the environment has less than 500000 polygon lines depending on their relevance in the context. Also we have reconstructed different artefacts.

We used for reconstruction 3D Studio Max and Maya applications converting each object in individual, identifiable and unique component in the environment. The objects are constructed for using in three different methods: frame interaction, 3D object interaction and dialog interaction. In Figure 2 is presented a section from inside Citadel reconstruction.

Figure 2. Landscape reconstruction of the Roman Temple object in the Roman Citadel Apulum

Textures are created using Adobe Photoshop based on the images taken on the Museum or in site. Some artefacts are reconstructed exactly like they appear in the exhibitions like coins, bowls, candelas, pottery, altar, grey ware beaker, statue, plate, hand made jars.

In Figure 3 is presented the UV mapping texture of the ancient coin, this object have 3D form and it will be used in the environment for real time interaction using haptic devices. Our future research will be oriented also to feel the surface of the coin, not just the dimension and position like in this application.

Figure 3. UV mapping of the coin texture

The texture dimension are memory scalable, that means that resolution for the picture are power off 2 (2^k) to increase the overload and memory speed storage. In that sense we take in to conssiderance a balance between dimension and detail or picture quality. We use also some free texture looking on the web sites (*www.cgtextures.com*).

2.3 Interaction with objects

In the section below we talk about three type of interaction with objects and environment similar with game interaction we propose. Because this application has an educational purpose we combine the game facilities with learning techniques and also interactive interfaces. All the interaction is based on the intersection between collisions mesh of the objects and collision ellipsoid of the camera.

Interaction with frames is based on the collision mesh integrated with The Irrlicht Graphic Engine. The Irrlicht Engine is an open source high performance real time 3D engine written and usable in C++ and also available for .NET languages. It is completely cross-platform, using 3D, OpenGL and its own software renderer, and has all of the state-of-the-art features which can be found in commercial 3D engines (Gebhardt N., 2007). Interactions are generated by the collisions of two or more than too collision mashes. We define camera to be the visitors, the player in the environment, ho has their own collision ellipsoid mash. The player can change the axis radius of the ellipsoid in real time using a short key on line. In that case a menu appear and player can change some futures of the environment like distances, light, visibility radius, frog or other parameters. When a collisions, ore more, appear the player may choose witch collision is interested for them and start the action specific to this selected collision.

Interaction with frame generate text or sound explication about the environmental or object interactions, storage in text or sound file each of them identified by unique ID. For each collision mash defined, also for frame, object or persons, is ID defined storage in the xml file(in the section data base and algorithms we explain the used technique).

In order to launch the event associate with the frame, object or persons interaction is defined a database trigger, a procedural code that is automatically executed in response to certain events on a particular table in database associate.

Interactions with object is similar like the frame interaction execute a trigger associate in two different ways. If the haptic devices are note connected to the computer a window appear on the screen containing 3D visualisation of the object guided by tree scroll line corresponding to rotation horizontal, vertical and scale object.

In case of haptic device the same window show the object evolution coordinate by the glove. We used in the experiments Pinch Glove system for interaction. Pinch Glove is a system for interacting with 3D simulation. This pair of stretch-gloves contains sensors in each fingertip which detect contact between the digits of your hand. It is used these gestures for a wide range of control and interactive functions customized to the specifications. Any combination of single or multiple contacts between two or more digits can be programmed to have specific meanings, ranging from simple on/off to multi-part, multi-action commands. The gestures are not dependent on individual hand geometry. The Pinch never requires calibration.

No complicated gesture recognition algorithms are required, so it is easy to integrate the Pinch into interactive 3D video games or any application that requires a wide range of tactile gestures. Interaction with persons is also mash collision based but if a person is occurred different trigger is launched.

Figure 4: Interaction with roman soldier using voice and text.

Interaction with persons is dialog based. In Figure 4 an object interaction occurs with person. A trigger is launched with possible questions, answers and written text or sound dialog. The sound dialog is based on two different forms. Interaction with frames generate an explanation form both text and sound without outside interaction. Interactions with person generate both text and sound interactive dialog based on the graph dialog stored in the database. For the beginning we have started with storage interaction dialog, but for the future we intend to develop an artificial intelligent system to upgrade the person knowledge in real time.

The dialogs contain educational questions and answers guided in to graph dialog based algorithm.

3. DATA STRUCTURES AND PROGRAMING

Our research is based on the simple and quickly data base access, open structure, free software modules, free engines and C# programming language application.

3.1 Data base structures

The data base structure is html oriented based, easy to update, universal recognized and program languages interpretable. We choose that way to organize data because classical data structure means servers or local data files access increasing time transfer and data search. The database is structured in different html file with two types of data, each of them heaving similar structure. The structure for objects are described in Formula (1) and explained below.

Example:

<Object type="menu"
id="1"
mesh="../final/moneda1.my3d"
mesh_viewer="../final/moneda1v.my3d" (1)
mesh_ viewer_scale= "2 2 2"
desc="Test text moneda 1." />

where *Object type* = define the objects type, that may be menu, person or different objects defined
id = is numbered identification of the objects
mesh = contain the path to identify the mesh file
mesh_viewer = contain the path to identify the mesh viewer file
mesh_ viewer_scale= contain the path to identify the mesh viewer scale in three dimension x y z, separate by spaces

We used a binary search algorithm in the special html file to look for different elements, identified by their *id*. The object types determine the interaction type. Collision between camera ellipsoid and object collision mesh determine specific action. The action is defined by the trigger type, which may be windowed information, sound player and person interaction or object viewer interaction. The trigger data base format is presented in Formula (2).

Example:

<Trigger type="switch"
state="1"
id="12"
desc="Cazarma fortificata" (2)
meta=""
visible="0"
position="-5 23 -861"
scale="2 2 2"/>

where *Trigger type* = define de trigger type
state = the value 0 or 1 define the state of the trigger, not interaction or interaction with the object
id, desc, scale = are similar with formula (1)

meta = meta information's specify the input or output line in/off dialog graph, running parallel tasks like sound and text
visible = some elements like windows, information or menus may be invisible marked with 0 or visible marked with 1.
position = mark the real position in the tree dimensions space starting with origin in the middle of the general frame, meaning that we have also positive and negative values.

The dialog with person is structured in the *hml* file, dialog line with dialog line, witch have a special format presented in Formula (3), where two examples for entrance and for exit are presented. Lines are numbered that means each of them is an entrance in graph dialog containing the list of possible outputs line or null for exit.

Examples: <Line meta=""
npc = "Hello!"
pc="Question?"
vec="1,2,3"/>
<Line meta="exit" (3)
npc = "Answer"
pc="Bye!"
vec=""/>

where Line meta = the same like in Formula (2)
npc = non playable characters, text answer to given question
pc = playable characters, possible questions
vec = a vector, containing separate by coma, possible answers line numbers, that follow the conversation starting with this question. Null is end of conversation.

Based on this data structure the player is passing question by question, answer by answer in the dialog graph. The question and answers are followed also by the sounds. The future development is to integrate the Microsoft speech but only in English. Those features are supported by .NET.

3.2 Program structure and algorithms

Application was created using Microsoft .Net C# language, object oriented platform. To integrate the open source modules and free IRRLICHT engines it was necessary to create three different classes.
In Figure 5 is shown the Engines Integration Diagram. IRRLICHT Graphic Engine is the main engine in relations with IRRNEWT Wraper, providing for IRRKLANG Sound Engine resources and control.

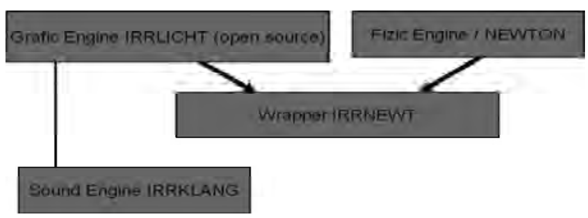


Figure 5: Engines integration diagram

Movement and motion is realized using Physic Engine Newton. The wrapper aim is to integrate objects in the frame, establish the mash for every environmental object and reconstruct the frame in real time.
We try to increase graphics loading, texture reconstruction, collisions conflicts and data base access to information or sound effects.
The application structured class based, containing *CApp*, main class, *CLevel*, designed for static geometry, objects management, interaction with objects and dialog with persons, *CEventReceiver*, for input output management.
We had create a main class, *CApp*, with the structure designed for load settings, load level controls, get objects ID's, control the player camera, animation, viewers, image loading, trigger control. To reduce memory dimensions of the images we use crypt and decrypt methods for image loading and saving. Interactions with the objects are determined using collision mesh, bound by the 3D *bounding BOX*, like in Figure 6. The algorithm determine the radius R between players camera and bounding box centre, calculating minimum distances collision necessary for interaction.

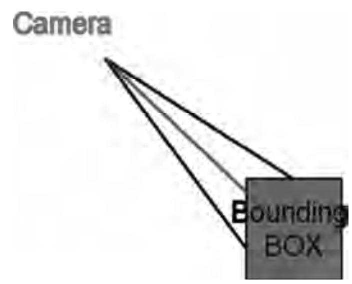


Figure 6: Camera radius R, calculated for object bounding box

CLevel class contain methods for static geometry engine methods accesses. *CLevel* class have methods designed for distance visibility and light flags.
CEventReceiver contain important methods for object triggers and associate event objects in input output devices. These methods describe and offer facility for chose different event in interactions with the objects. Interfaces are created using special windows for presentation, object interaction in view mode or interaction mode.
Important facilities developed in the application are: interactive menu witch offer possibility, for player, to change in time some interfaces characteristics, light effects, viewed distances, fog effects offered by Engine.

The main algorithm tries to maximize the efficiency loading scenes only if the player is on the site. Initialization starts with the nearest objects from the camera loading. The mesh object on the area is loaded in scene keeping in memory just coordinates of the objects identified by their ID's. The camera position determine the list of the loaded objects, they are loaded by the graphic and physic engine and shown. The object structure is defined by their components: Id, name, description, type, relative position in 3D representation, mesh dimension for collisions in three directions, visibility and scene address.
Using the camera coordinates and objects coordinates the algorithm loop determines in real time collisions between camera and other objects and starts the corresponding trigger.

Corresponding with started trigger some windowed interfaces are displayed depending on the object interaction. The player have the possibility to interact with the object in the new window created or come back to the scene to look for another interaction. Shot keys are defined for easier interaction and rapidly interface parameters to be changed.
All the features of the application are growing up while the project is develop and we hope that interfaces are useful and easy to use for all the players.

4. CONCLUSIONS

Historical information was obtained with the researcher support from National Museum Alba Iulia, Ph. D. Moga Valer and Ph. D candidate Lascu Ilie. We mention that that research is financed by the "The National Cultural Found Administration", in two projects "Muzeul Virtual 3D", (3D Virtual Museum) in 2007 (*www.afcn.ro/beneficiari_castigatori_sesiune_noiembrie _2005.php*) and "Traseul celor trei Fortificaþii" (Three Fortress Ride) in 2008 (*www.afcn.ro/finantari_proiecte_culturale*). The second project is under development until October 2008, some reconstruction will be made in the next months and it will be included in application.
The research is oriented to look for the simply end easiest ways to develop interactive 3D presentations used in the Cultural Heritage promotions. The combination between free software engines, interactive application and scientific archaeological documentation create a double effect: a good documentation and scientific support for future development in archaeological research, providing a learning interactive scene suitable for kids and heritage interactive presentations. The aims of this research are to obtain a real time interaction on the web using haptic devices. This platform is suitable in that sense. With the support of the archaeological team we try to develop a Real Time Virtual Platform for Heritage visitors on the web.

5. REFERENCES AND SELECTED BIBLIOGRAPHY

References from Journals:

Anghel Gh., AIIA, Anuarul Institutului de Istorie ºi Arheologie, Cluj (Napoca), XIII, 1985, p. 264.

References from Books:

Barcelo J., Forte M., Sanders D., 2000 (eds.), *Virtual reality in archaeology*, Oxford, ArcheoPress (BAR International Series S 843), pp. 247-263.

Moga V., *Castrul roman de la Apulum*, Editura Casa Cãrþii de ªtiinþã, Cluj, 1998, în SIB, XXIII-XXIV-XXV (1999-2001), 2002, Timiºoara, p. 44 - 70.

References from Other Literature:

Anghel Gh., *Noi date în legãturã cu vechea cetate medievalã de la Alba Iulia*, Apulum, XXII, 1985, p. 113-122.

Camporesi C., *Generazione di paesaggi tridimensionali. Un approccio OpenSource per applicazioni real-time e web-based*, dissertation of univ. of Bologna, Dept. of Computer Science (Prof. G. Casciola) Calori L., Guidazzoli A., Forte M., 3D Virtual scientific visualization of archaeological sites: new navigation metaphores, in *Scienza e Supercalcolo*, CINECA luglio 1996, Bologna, pp.9-12.

Calori L., Diamanti T., Guidazzoli A., Liguori M.C., Mauri M.A., and Valentini L. (2003), *Certosa virtual museum: a dynamic multilevel desktop VR application*, poster in Eurographics 2003 Proceedings, 1-6 September 2003, Granada, Spain.

Kovacs A., *Construcþii patronate de Gabriel Bethlen*. Ph. D. Theza, Cluj Napoca, 1984, p. 45.

Moga V., *Apulum, porta principalis dextra a castrului legiunii XIII Geminae*, National MuzeumAlba Iulia, 1999, p. 12.

References from websites:

Calori L., Forte M., Pescarin S., *Real-time interactive reconstruction of archaeological Landscapes: an opensource approach. From GIS to Virtual Reality*, CINECA, Italy, 2003, http://3d.cineca.it/3d/pubblicazioni/2003/Italy-Canada05Workshop_pescarin.pdf (accessed 2 dec. 2007)

Gebhardt N., *Irrlicht lightning fast real time 3D engine*, ghttp://irrlicht.sourceforge.net/ (accessed 10 oct. 2007)

www.cgtextures.com (accessed 12 aug. 2007)

www.afcn.ro/finantari_proiecte_culturale.php, *Patrimoniu cultural national la 30 aprilie 2008.xls* (accessed 20 June 2008)

www.afcn.ro/beneficiari_castigatori_sesiune_noiembrie_2005.php, *Catalogul proiectelor culturale finantate in anul 2007.xls* (accessed 20 June 2008)

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

VIEW WHAT YOU WANT HOW YOU WANT: COMBINING DATABASE SYSTEMS AND CUSTOMIZABLE VIRTUAL REALITY TECHNIQUES IN CULTURAL CONTEXT

S. Valtolina

Dept. of Computer Science and Communication, University "degli Studi" of Milano, via Comelico 39, Milano, Italy - valtolin@dico.unimi.it

KEY WORDS: Virtual Museum; Adaptable Content; Knowledge Management

ABSTRACT:

The innovations promoted by the e-society highlight the central role that the e-culture has to play as a privileged means for disseminating cultural heritage. In this area, one weakness affecting exhibitions held in conventional museums is that generally the artefacts are separated from their cultural and historical context. Connections among artefacts belonging to collections owned by different cultural institutions are also difficult to display. Virtual reality (VR) reconstructions can be useful in recovering an artefact's context also through web sites. However, virtual worlds usually offer little opportunity for personalization. This paper presents Ci.V.E.D.I+, a system developed to address the above issues. Ci.V.E.D.I+ integrates a VR environment generator engine accessible via web, customizable according to the user preferences, with an interface to access contents from digital archives. The re-contextualization of the cultural heritage is carried out using narrations, that is, stories written by domain experts about one or multiple artefacts. Ci.V.E.D.I+ dynamically retrieves from the digital archives the elements referenced by a narration and inserts them in the VR environment.

1. INTRODUCTION

1.1 Virtual Reality and Cultural Heritage

The goal of virtual reality (VR) is to recreate, analyze, explore 3D objects and worlds in virtual environments and to stimulate the main human senses, mainly the sight, but also the hearing and the touch. According to this definition, VR techniques can be applied to a large variety of domains in which they have a significant potential for supporting the visualisation of 3D materials, images, sounds, datasets through a comprehensive and visually appealing presentations also through web sites.

As such VR has an important potential for enhancing many aspects of the human life in particular in the field of the cultural heritage valorisation and dissemination (White, 2004). In particular, because VR overcomes time and space barriers, it offers the possibility to virtually access environments that are no longer accessible, to visit museums located far from the user, and to come in contact with cultures in forms which are not possible in real life (Loscos, 2004). Again, VR offers a space in which users can meet other people, exchange experiences, learn, and interact with domain experts such as archaeologists, historians or art critics. In this respect, the combination of VR and technologies such as the Internet is the key to make cultural heritage available anytime and anywhere. However, in order for this process to happen and the large public to benefit from it, it is necessary to build systems having a high communicative impact and to assure an equally high scientific quality of the cultural contents. In this way exhibitions in VR can offer a new method to capture the museums information allowing their contents to be better explored and to better match the expectation of the users.

However, to date, most of their solutions fall into the typical windowing, 2D paradigm, thus missing the benefits of the attractive nature of 3D especially for a public of non expert users. Think as examples of web sites presenting paintings: a distracted user might not realize that Leonardo da Vinci's Monalisa ("virtual exhibited" at the Louvre museum website) is more than sixty six times smaller than Picasso's Guernica ("virtual exhibited" at the website of the museum of Reina Sofia in Madrid). Of course the size of the painted canvas is not the most important element of a painting, but systems promoting cultural heritage to large public cannot ignore visual elements such as the size of the artefact, especially when multiple elements are presented together.

These kind of issues are very current problems since today several museums and cultural institutes integrate their expositive heritage using virtual environments accessible by web or by other installations placed in real spaces (White, 2004). Through these technologies, users can explore the museum exhibits in a VR that is both spatially accurate and visually engaging. Besides, during this kind of visits, surfers find in the virtual space a familiar environment able to adapt to their expectations and visiting styles. However, those systems are expensive to build (especially for small museums) and, more important, they do not usually support enough interaction and customization for users who, in most cases, either have to follow a predefined tour or can choose only among a limited set of predefined options. To overcome these difficulties it is important to develop systems supporting the dynamic creation of virtual worlds in which users are able to browse the cultural domain according to their needs and wishes. A previous paper (Mazzoleni, 2004) has tackled this problem developing a system, referred to as Ci.V.E.D.I (Customized Virtual Environment for Database Interaction), in order to create flexible and personalized VR galleries, accessible via web, in which works of arts are dynamically loaded in the virtual environment, by querying a Multimedia Database (MMDB). By using Ci.V.E.D.I, the user can not only specify the content of the virtual gallery, but also customize the environment in which the exhibition is arranged, the organization of the artefacts within the visit, and the overall duration of the visit itself. Ci.V.E.D.I separates the content (extracted from a relational database) from the virtual environment that becomes a sort of gateway used to obtain information of interest to the user. Ci.V.E.D.I represents a viable solution for small museums interested in exhibiting their collections in a VR environment in which the visitor is engaged in a customized tour that does not necessarily reflect the ones possible in the physical world.

However, Ci.V.E.D.I is not enough to promote the VR as a channel to favour the mutual exchange of cultural heritages. Ci.V.E.D.I lacks essential features for supporting the users in managing, manipulating, distributing and sharing knowledge.
In other words, one weakness affecting this digital reconstruction of exhibitions is that the artefacts are separated from their cultural and historical context. Two main aspects need to be considered. First, virtual museums should not be built only using "virtual artefacts" retrieved from single collection, but different museums should collaborate exchanging their collections as it happens today for temporary exhibitions. Second, databases, in many cases, do not provide mechanisms supporting the description of the anthropological, historical, artistic connections among the artefacts. In real life, the museum curator decides which "theme" to represent in the exhibition and how to organize the artefacts in the expositive areas. Domain experts (as well as all people who think to have something to say about a set of artefacts) should be able to participate by specifying their own contributions and allowing the visitors to choose the "thematic" tour(s) which better fit their interests.

This paper tackles such problems describing how to extend the parameterized environment built using Ci.V.E.D.I in order to maintain the semantic correlations among the elements retrieved from digital archives of different museums, and to augment them with contextual information. This extension, called Ci.V.E.D.I+, allow to integrate a VR environment generator engine, customizable according to the user preferences, with an interface to access contents from digital archives. Thus, in Ci.V.E.D.I+, the information is dynamically retrieved from multiple data sources and museums can decide to participate with their artefacts to a virtual exhibition at anytime and without any need of replicating their data. Under this perspective, Ci.V.E.D.I+ presents an innovative prototype for the creation, aggregation and dissemination of knowledge making it sustainable and visible in VR environments accessible via web. The re-contextualization of the cultural heritage is carried out using narrations, that is, texts, short essays, written by domain experts, useful to contextualize the single artefacts retrieved from different archives, through their connection with specific topics and spheres of knowledge. By means of the narrations it is possible to tailor the cultural heritage information according to specifications defined by domain experts. In this way the user can properly navigate through the heritage and create their own personalized thematic tour through a large number of information trails. Ci.V.E.D.I+ dynamically retrieves from the digital archives the elements referenced by a narration and inserts them in the VR environment

Essentially, the contribution of the paper is twofold. First, it recognizes VR as a privileged channel to exchange cultural heritage which can overcome the limitations of the 2D windowing interface today adopted by many museums' website. Second, it presents Ci.V.E.D.I+, a system to dynamically build virtual galleries in which museums, domain experts, and visitors collaborate in deciding the content to be presented according to their knowledge and interests. With Ci.V.E.D.I+, the virtual environment becomes a single point of access to the heritage distributed in multiple data sources contextualized by the contribution of autonomous domain experts.

The rest of the paper is organized as follows. Section 2 compares different approaches to develop user driven VR environments. Section 3 presents the approach to the development of the Ci.V.E.D.I system and how it has been extended in the Ci.V.E.D.I+ system. Finally Section 4 concludes the paper and outlines future works.

2. RELATED WORKS

2.1 VR Techniques in Cultural Field

During the last decades, a large number of projects related to the application of VR technologies to cultural heritage have been developed. 3D scanning, modelling and rendering techniques have been used for the reconstruction of historical artefacts (Jablonka, 2002), whereas virtual and augmented realty environments have made available new interaction models to users and experts (White, 2004; Papagiannakis, 2004; Sinclair, 2005; Azuma, 2001). Archaeology is a typical case, in which the virtual cultural heritage can communicate complex information according to what is actually visible and what is not, contextualizing the real and the imaginary for a correct interpretation of the present and the past. Virtual archaeology allows one to use scientific data for recreating monuments, findings and environments that the time has turned into fragmented ruins. Examples of the application of computer graphic, multimedia, robotic and VR to this field were presented during the world exposition on the ancient Rome in VR from September 15th to November 15th 2005 in Rome (CNR-ITABC, 2005). Despite their high quality, these applications must constantly match with the tremendously high expectations of the users, created by illusions of what VR could be provided according to what movies, television, and literature state. For example J. Lanier (Lanier, 1992) discusses motivations why VR has not yet become a widespread technology. The main reasons concern: low computer performance, too low resolution screens, no clear and complete understanding of the human cognition, acuity of human senses, sense of isolation from the real world. The performance lacks will be overcome by the growing computation power of new processors and by innovative hardware technologies, whereas issues connected to human perceptions will be tackled by new interaction models supporting content personalization according the users' wishes and needs. In fact the user's expectation is high because she/he is not attracted by simple walk through application but she/he wants to live a formative experience and to focus on her/his topics of interest. Several projects have investigated different intuitive and customizable access to cultural and historical information retrieved from museums and archaeological sites using mobile devices, multimodal interfaces or other interaction devices able to make the scene less inert and users more active (Ciger, 2003; Valtolina, 2005; Champion, 2004). However the interaction dynamics, at the base of the customization of the user's visit, has not been much investigated. Applications should encourage learning and personal enrichment. In order to achieve this goal, information should be shown according the user's background, and the artefacts should be arranged in the virtual set according to their natural and semantic correlations. The overall experience can be optimized by enriching the exhibition with tours presented to the visitor in a coherent way and fixed according to contributions of expert of specific domains. To show a virtual environment, created starting from specific user's requests, it is necessary to exploit a wide base of knowledge from which to retrieve information generally owned by different museums and cultural institutions. Several projects have faced this problem (Ammoura, 2001; Costabile, 1998). For example, the main goal of the ARCO project (White, 2004) is the development of technologies to create virtual representation of museum's artefacts using a stereo augmented reality system. Digital photographs, 3D models and other descriptive metadata are organized in an object oriented database used to create virtual exhibitions. These representations are then visualized according to different modalities through a VRML or X3D (Brutzman, 2007) based web browser or directly in the museum using Augmented Reality Interfaces. Another interesting project is Sculpteur (Sinclair, 2005). In Sculpteur, 3D objects retrieved from the database by using specialized retrieval algorithms are associated with concepts defined in a special purpose semantic layer. By mean of a "concept browser" users can access to the information from the partner museum's through a common ontology, the CIDOC CRM (Crofts, 2008). In addition, classifier agents can be trained to make automatic classifications of objects in order to present the final virtual exhibition. However in these projects only a single database is used, which thus makes difficult to make available information from several museum's; furthermore, the 3D artefacts are arranged according to some simple classification criteria, thus without representing the semantic correlations among them.

In another project, art-E-fact (Marcos, 2005), one of the goals is to develop a generic platform for interactive storytelling in Mixed Reality that allows artists and content generators to create artistic expressions in a cultural context between the virtual and the physical reality. By means of an Authoring Tool, artists, users and content generators are able to create stories based on the content stored in databases.

The approach presented in this paper provides the additional information useful to contextualize the artefact through narrations, which are stories created by domain's experts. Narrations are integrated in virtual exhibitions thus enriching the user's experience and offering tours dynamically generated according the user's preferences. In order to achieve this goal digital images, 3D objects and related metadata are retrieved from databases owned by different museums that collaborate in the creation of a semantic layer representing the cultural domain of interest.

3. CI.V.E.D.I. APPROACH

This section describes the architecture of the Ci.V.E.D.I.+ system and how it has been developed from the Ci.V.E.D.I system. The goal of Ci.V.E.D.I. is to generate personalized virtual galleries of artefacts whereas Ci.V.E.D.I.+ proposes an interesting solution in terms of flexibility and extensibility. In particular it defines a method to overcome the limitation of extracting data from a single database and it offers the possibility for domain experts to share their knowledge by both enriching and guiding the visitors while accessing the virtual galleries.

3.1 Customized VR Environments

A main limitation of the current VR systems is that the information content is predefined and the user has a limited possibilities to change, customize or filter the information content according to her/his preferences, in other words there is not a suitable support for a correct arrangement and dissemination of the knowledge. Instead, in a customizable environment the user can choose the portion of information content by querying the system in order to present only the knowledge of interest. Ci.V.E.D.I. is a scalable system whose main goal is to provide a flexible and personalized virtual environment for accessing multimedia contents. By using Ci.V.E.D.I., the final user not only can specify the information she/he is interested in, but she/he can fully customize the appearance and the duration of the visit. User preferences are also integrated with the requirements from the curator of the exhibition, who, on the basis of her/his knowledge about the contents, can provide the system with additional information in order to generate tours that better facilitate the fruition of the available cultural heritage.

Starting from Ci.V.E.D.I architecture, Ci.V.E.D.I.+ proposes an approach based on the combination of multimedia databases and VR or, more precisely, the exploitation of VR as a solution to retrieve information from multimedia databases and to present them. The integration of VR and database technologies entails several issues, mainly related to the fact that the types and number of objects returned by a user query cannot be know in advance. Thus, the VR system must be able to dynamically rearrange the VR assignment, in order to be able to accommodate query results.

The methodology adopted to design and build customized virtual environments for delivering multimedia contents can be divided into two main phases: the "Data specification phase" and the "Tour generation phase" (Figure 1).

In the "Data Specification Phase" (Figure 1), the responsible of the cultural institution (i.e., the museum curator) provides the system with all the information needed to generate the tours. In this phase, the curator defines the subset of the data stored in the database that can be accessed by the user. After this, the curator selects the morphologies of the museums in which the exhibition can be arranged. In this solution, the exhibition space is not fixed, but it dynamically generated from a set of rooms each one defined according to a specific architectural style. In other words, a virtual room defined by means of a 3D template, called Virtual Scene Unit (VSU), which specifies the room's planimetry and its appearance (e.g., the textures applied on the geometry, the number and position of light sources, the intensity of each lights, etc). In addition, VSU contains other information concerning the expositive areas, which are the available zone on the room's walls in which it is possible to arrange multimedia elements, and additional parameters the system follows when it needs to resize the virtual environment, in case the total amount of space required by the paintings to be exposed exceeds the room's expositive area. The curator can choose among the available room templates, or can design its own VSUs, in order to define her/his personalized style (e.g., a modern museum, a church, a "Galleria degli Uffizi"-like structure, and so forth) which is more appropriate to exhibit the heritage in the database. The number and the size of the virtual rooms, the position of each single painting, the position and the size of the doors connecting two rooms, as well as the overall organization of the museum's rooms is dynamically computed based on the information to be presented.

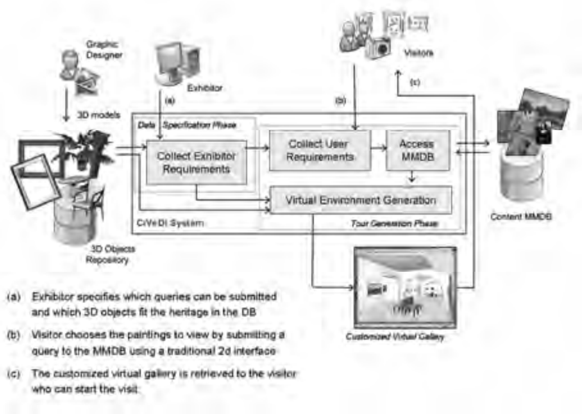


Figure 1: Ci.V.E.D.I+ system architecture.

The "Tour Generation Phase" (Figure 1) aims at generating the customized virtual tour according to user requirements. Using a traditional 2D interface, the user specifies her/his preferences, choosing among the alternatives made available by the museum curator during the "Data Specification Phase". She/he can select the query she/he wishes to submit to the database (e.g., returns all Picasso's paintings), the criterion to group the retrieved paintings (e.g. the place in which they have been painted, the year of creation, etc.), and the museum's architecture in which to arrange the visit. Once all preferences have been collected, the system constructs the necessary query statement.

During this step, the system generates a room among the VSU with the architecture selected by the user. Before being placed in the virtual room, the images are automatically resized with respect of the proportion of their real counterparts. Paintings are clustered according to the organization criterion selected by the visitor and each object in the cluster assigned to one of the expositive area available in the virtual room. Every time a painting is allocated, its area is subtracted from the total amount of free space in the expositive area. If all the objects in the cluster do not fit in the selected VSU, some relaxation techniques such as scene enlargement (within the bounds fixed by the VSU) or the organization the artefacts in two parallel lines instead of one in each wall are applied. Furthermore, there will be cases in which the system will load another scene until all the objects have been assigned to a VSU.

As a final task, before applying the algorithms to draw 3D scenes and objects composing the user customized virtual exhibition, Ci.V.E.D.I. applies additional heuristics to refine the results of the previous phases. Specifically, the system reduces VSU dimensions in cases artefacts cover only a small percentage of the available expositive area. If it is not possible to reduce the VSU, the free expositive areas left with no paintings are automatically filled by autonomous objects such as windows or columns, etc.

3.2 An Integrated Access to the Cultural Heritage

To provide an environment in which VR is a channel supporting the mutual exchange of cultural heritage, two main aspects need to be considered. First, virtual museums should not be built only using "virtual artefacts" retrieved from single collection, but different museums should collaborate exchanging their collections. Second, the "virtual artefacts" have to be presented according to the their anthropological, historical, artistic meaning and the semantic relationships among the visualized artefacts have to be highlighted.

Starting from these considerations and exploiting the architecture defined in the Ci.V.E.D.I system, the Ci.V.E.D.I.+ system defines how to maintain the semantic correlations among the elements retrieved from digital archives of different museums, and to augment them with contextual information. In Ci.V.E.D.I+, the information is dynamically retrieved from multiple data sources and museums can decide to participate with their artefacts to a virtual exhibition at anytime and without any need of replicating their data.

The integration among the various information sources is achieved by means of a semantic model, represented by an ontology (Valtolina, 2007) and able to describe the relevant concepts of the cultural heritage domain. In this context, VR technology is used to create environments that, by hiding the complexity of the underlying databases and archives, are more natural to users and easier to use when interacting with the cultural heritage information.

Although many institutions and museums have well organized digital archives, these communities still lack effective and efficient technological supports for collaborative knowledge networks. The most outstanding innovative feature of the such a network is its ability to bring together information, which is normally distant, because of geographical constraints and/or institutional background. In the data integration context a lot of proposals have been introduced, (Bergamaschi, 2001; Leone, 2005) and basing on these studies, it is possible to affirm that the use of an ontologies for the explication of the common knowledge seems to be a promising method for addressing the integration problem. In particular the integration model adopted in the Ci.V.E.D.I.+ system , is based on a use of an ontology as knowledge base representation where data residing at the sources are accessed during query execution. This approach called virtual approach (Lenzerini, 2002; May, 2005) uses the ontology as a sort of semantic access point of the information that can be retrieved from databases. Databases owned by different institutions are mapped and related independently from their number, location and typology. In this way, the information are not recreated, but extracted automatically from the existing sources through the ontology schema. Nevertheless, creating a new ontology from scratch is a time consuming task and, therefore, it is better to exploit a general ontology from which customized cultural ontologies can be derived. In the proposed solution this ontology has been derived from a well known cultural heritage ontology, that is, the CIDOC Conceptual Reference Model (CRM) (Crofts, 2008). The CIDOC-CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation and it is a well formalized, well supported and widely accepted standard in the cultural heritage area. Once defined the suitable ontology for representing the cultural information domain, it is necessary to establish a correspondence between the database schemas and the ontology classes and properties. In this way, the system can determine how the ontology classes are represented in the databases and how to access them. The result of this process, iterated for each integrated database, is a semantic network able to translate the database schemas onto classes and properties of the ontology. The information defining the mappings are used by the final system to generate SQL code for querying each database schema. In this way the query can be expressed using semantic queries independently from the query languages of the underlying databases.

Besides an integration problem of the information domain, nowadays there is also a problem of dissemination of the information. Several times in the museums the artefacts are decontextualized and are often juxtaposed to other objects on the base of properties such as chronology, typology, taxonomy, shape, stylistic schools and so on. This kind of exhibitions emphasizes the formal aspects of objects rather than their thematic contents which are useful to promote a better understanding of the life, the behaviours, the culture of ancient people, in relation with cultural models of the past and of the present. Therefore an innovative notion of cultural heritage virtual display would see the museum as a virtual space where visitors could find, close to each other, and compared, physically distant objects, texts and artefacts from different countries but belonging to similar cultural horizons (history, style, society).

In order to customize and contextualize the user's visit according to the historical, artistic, or anthropological meanings of the displayed artefacts, it is necessary a direct participation of the domain experts in the construction process of the final applications. Only domain experts are able to enhance the information presented in a virtual environment by guiding the attention of the end users to the elements which are more relevant to a specific theme. This is what happens in a real visit: the museum curator arranges the artefacts in the available space according to her/his knowledge, thus defining a specific "theme" for the exhibition; she/he eventually provides additional information to enable the visitor to better understand the suggested theme. In the virtual world, the goal of automatically generating such thematic exhibitions is much more challenging, because there is a large number of works of arts that can be presented and a lot of domain experts available to specify their "themes" or "themes of themes". However, the virtual exhibition should not be managed according to the indication of a single "curator"; instead, domain's experts (researchers, museums curators, art critics, painters, etc) should all be able to participate in the process by both adding their anthropological, historical, and artistic knowledge about the artefacts and deciding the content and the organization of the virtual gallery. Consider as an example the set up of a virtual exhibition of paintings produced during the XIX century. A domain expert might want to describe the biographies of the artists and to organize the paintings by artist. Another expert might want to describe the various phases of the cubism movement and its main representatives. In the same direction, other domain experts might want to describe how the use of perspective in paint has evolved over time and how Picasso introduced "Time" as a new dimension in his works of arts. All such information is not typically part of the database; however it is invaluable in helping users to enhance their knowledge about the works of art, and in suggesting new interpretations for them.

To address such issue, Ci.V.E.D.I.+ provides a methodology, based on the concept of narration through which domain experts can contextualize and connect different artefacts in order to generate customized presentations of the contents. A narration is a text, short essays, written by domain experts, useful to contextualize the single artefacts retrieved from different archives, through its connections with specific topics and spheres of knowledge. By means of the narrations it is possible to tailor the cultural information according to specifications defined by domain experts. In this way the user can properly navigate through the heritage and create their own personalized thematic tour through a large number of information trails.

To integrate the narration concept into the Ci.V.E.D.I.+ system, it is necessary to extend the ontology's schema by means of a new class called "Narration" and to enrich the "Data specification phase" in order to allow the user to select elements which better meet her/his interests. As described in the previous section, during the "Data specification phase" the user can decide to build her/his virtual tour selecting the paintings of interest according some criteria. Now the "Data specification phase" is enriched with the possibility for the user to browse a list of all the available narrations and select one of them. Afterwards, the user selects the environment in which data should be organized. The information (narrations and paintings) that the Ci.V.E.D.I.+ system determines as relevant with respect to the requirements specified by the visitor is then retrieved and given in input to the "Tour Generation Phase", in which the system will create the virtual environment.

It is important to remark that each narration is a full fledged, self contained story, which can be understood independently from any other external information. However, different narrations can discuss related topics, and so it is possible to link them. For example, a narration about "The evolution of painting during the XIX century" provides some background information to the narrations "The Cubism" and "The Futurism" concerning the two artistic movements; instead, a temporal relation can be established between the latter two narrations (Futurism is subsequent to Cubism).

These relations implicitly express a network of semantic links connecting the narrations' subjects. It is possible to model these relations by a set of properties specifically introduced in the semantic model; for example, a property is used to organize the narration according to a hierarchical structure, in which subjects of the topmost narrations are more general than that of the lower ones. In the previous example, the narrations on Cubism and Futurism have a topic which can be referred to the context of the narration "The evolution of painting during the XIX century". Other properties describe temporal relations, like the one existing between the narrations on Cubism and Futurism, or express the role that a narration plays towards another narration, i.e. narrations concerning Picasso and G. Braque "takes part to" narrations concerning the Cubism movement.

A narration linked to another narration through one of the properties previously discussed is called "narration_element". "Narration_elements" are taken into account by the system during the virtual environment generation process. More precisely, after creating a virtual room for each narration selected by the user, the logic of the property linking narrations and "narration_element" is evaluated to create a continuous structure. For instance, consider the three narrations introduced above. In this scenario, the narration "The evolution of painting during the XIX century", which provides the context for all the other narrations, will be placed in the centre of the exhibition area; the user will start the visit from there. The two narrations "The Cubism" and "The Futurism" will be organized in two subsequent rooms, whereas the sub narrations "Picasso" and "G. Braque" will be organized in two rooms both connected to the room associated with the narration "The Cubism". Sets of paintings referred by different "narration_element" can be merged in the same area if their cardinalities, taken individually, are too small to require a new room for the paintings allocation

It must be emphasized that all those operations are automatically executed by the system without any human intervention. With such approach, the virtual museum architecture is not a “sterile” environment in which paintings are arranged; instead the architecture of the virtual rooms contributes to present the heritage according to a specific theme. The resulting virtual gallery is then presented to the visitor who can start the visit on her/his personalized museum.

3.3 The Ci.V.E.D.I.+ Prototype

A prototype of Ci.V.E.D.I.+ has been implemented using Java as programming language and in accessible via web. The virtual environments are defined with X3D (Brutzman, 2007). Virtual worlds created by X3D can be presented via a common web browser and integrated with other web based applications. The combination Java&X3D has been chosen to reuse the Ci.V.E.D.I.+ components which arrange the paintings in the VSU’s and which draw the final world. A difficult issue in the development of the system is related to the presentation of the textual information (narrations and details about the exposed paintings) in the virtual gallery. In this context, it is possible to automatically organize in the virtual galleries a net of active areas, invisible to the user, which are associated with either a narration or a painting. When the user crosses an active area associated with a narration, the narration’s text is shown to the user. It is not rendered in the virtual environment, which is typically not suited to handle textual information; instead, it is displayed in two dimensional boxes manageable as traditional windows. As an alternative, the user can choose to hear a pre-recorded audio stream in which the narration’s text is read by a human speaker.

Again, when the visitor gets closer to a painting, the active area triggers a request to the Ci.V.E.D.I.+ system which retrieves from the database the additional information of the painting and all the narrations related to it. In this way, the user is free to leave the current narration to start another one. In this case, a new virtual scenario is generated and the visit starts from the last virtual room/painting viewed. In addition, the previous galleries are cached in case the user wants to return to her/his previous selection. Figure 2 shows an example of the virtual gallery created using Ci.V.E.D.I.+; in this case, both the paintings and the narration are presented to the visitor. The menu bar on the bottom of the screen is used to manage (play, stop, pause) the audio file associated with the narration. The last visited gallery is reachable through a hyperlink located in the bottom right corner of the screen. On the top right corner is a map is visible showing the user current position within the virtual environment.

Figure 2: Ci.V.E.D.I+ virtual gallery

Particular attention has been devoted, during both the design and implementation of the Ci.V.E.D.I.+ prototype, to usability requirements, which are clearly a key aspect for the success of the system. Following the ISO 9241 standard, three main usability indicators were adopted: effectiveness in use, efficiency in use, and satisfaction in use. At the present, in the context of the degree course in Digital Communication of the University of Milano, several students are carrying out several usability analysis exploiting the Heuristic evaluation methodology.

Heuristic evaluation (Dix, 2004) is a usability engineering method for finding the usability problems in an interactive system. A team of one or more HCI experts evaluates the application, mimicking the users according to a given profile. In the present case, several students are covering the role of evaluators and once finished their analysis the prototype will be redefined and improved according to the students’ feedbacks.

4. CONCLUSIONS AND FUTURE WORK

This paper has presented Ci.V.E.D.I+, a system to generate customized virtual galleries in which paintings are dynamically extracted from several multimedia databases. The system does not only offer the possibility for the visitor to enjoy a 3D visit of a museum in which the paintings are the ones which better fit her/his interests, but it also gives the unique opportunity to learn more about the paintings by accessing the narrations specified by domain experts. Domain experts are able to share their knowledge with others users (both visitors and other domain experts) by both adding content to the data in the digital archives and establishing organization criteria which are not explicitly represented in the database. On the other hand, museums can decide to join and leave the content network (database and narrations) at anytime and having their data represented in a 3D environment without having to move their content out from their databases. Ci.V.E.D.I+ is thus a comprehensive solution allowing all subjects involved in a cultural heritage domain to use VR to share their information with others with the goal of creating a cultural community of users.

Besides an extensive test of the developed prototype system, it is possible to suppose and plan several extensions as future works. First, a fist extension concerns the possibility for the visitor to actively participate in the creation of the knowledge linked to the cultural heritage domain. While visiting the virtual world, the visitor should be able to annotate, with her/his comments, both the exhibited paintings and the presented narrations. To know more of annotations please refer to (Fogli, 2004). Moreover the system should first generate all the environments corresponding to each narration, and then rearrange them in a single structure. This option is very important especially in case a user (for example a high school professor) wants to make available ad-hoc content for a restricted set of users (e.g., a class of students), but also to play the visitor's role.

Special efforts are devoted in studying techniques to let the system's users interact with each other. Tools like textual/vocal internet relay chat channel can be easily added as external component of the main application. However, the integration of this functionality inside the virtual tour is not a trivial task. In fact, this would require the development of a modified version of the X3D player plug-in supporting the messaging infrastructure. On the other hand, the integrated communication system would be the foundation for a whole set of collaborative tasks, like quiz games or treasure hunts, playable in the virtual environment.

Another possible extension concerns the possibility to integrate a text-to-speech engine in the system. This would free the content providers from the burden of creating the audio files associated with the narration and changing them whenever the narration's text is modified. As a last improvement regards the possibility to make the virtual environment generator engine able to handle not only bi-dimensional items, but also three-dimensional ones. Doing this, it would be possible to apply Ci.V.E.D.I+ to virtual heritage different from paintings, like sculptures, potteries, books, archaeological findings, etc., thus extending the system's applicative domain.

However, the solution proposed in Ci.V.E.D.I+ is built to be flexibly applied to many domains, even outside of the heritage scenarios. This because, nowadays, the use of new technologies such as VR is of increasing interest in several fields of human society not only the cultural field. In a large variety of domains VR techniques are applied in order to support the visualization and the analysis of 3D artefacts through a more comprehensive and visually appealing presentation. Exploiting the Ci.V.E.D.I+ system, current VR environments could be able to support an efficient knowledge transfer process. Appling the strategies defined in Ci.V.E.D.I+ it is possible to design virtual environments in which to offer a knowledge complete, accurate and consistently available to those who need it, when they need it.

At the moment a second prototype is under study and design. This prototype concerns the study and the design of a virtual environment called PUODARSI (Product User-Oriented Development based on Augmented Reality and interactive Simulation) in which to analyze, improve and optimize the visualization of 3D artefacts in a contextual and cooperative way. The system allows to reconstruct a virtual prototype in an augmented reality environment, to modify its shape using haptic devices and to assess the effects that these variations of shape have produced on the structural properties (by means of FEM - Finite Element Method) and fluid-dynamic properties (by means of CFD - Computational Fluid Dynamic analysis) of the object. In a virtual environment similar to which defined in Ci.V.E.D.I+, different experts, sharing their competences through the use of annotations, a different kind of narrations, exploit a unique simulation environment to design, modify virtual prototypes and execute scientific analysis in a collaborative and participative way. This new way of accessing, presenting and interacting with information content allows to change the user's role from the passive role of an observer to the active role of an knowledge worker. This approach contributes to the radical changes and innovation in the development from today's single application to globally shared cooperative spaces for a specific community. In particular PUODARSI addresses these issues in the context of virtual prototype design in industrial area. Experts belonging to different communities (designers, FEM engineers and CFD engineers) can interact simultaneously with virtual content displayed stimulating communication and encouraging joint discovery and problem solving process with distant users.

ACKNOWLEDGMENTS

Current researches related to the PUODARSI (Product User Oriented Development based on Augmented Reality and interactive SImulation) system, are partially funded by the Italian Ministry of University and Research (http://www.kaemart.it/puodarsi).

REFERENCES

Ammoura, A., 2001. DIVE-ON: From Databases to Virtual Reality. *The ACM Crossroads Database Issue*, Vol. 7, N. 3, pp. 4.

Azuma, R., et al., 2001. Recent advances in augmented reality. *Computer Graphics and Applications*, IEEE Computer Society Press, Vol. 21, N. 6, Nov/Dec 2001, pp. 34-47.

Bergamaschi, S., et al., 2001. Semantic integration of heterogeneous information sources. *In Data Knowledge Engineering* , Vol. 36, N. 3, 2001, pp 215-249.

Brutzman, D., et al., 2007 X3D: 3D Graphics for Web Authors. Morgan Kaufmann Publishers, Elsevier, 384 pages.

Champion, E., 2004. Heritage Role Playing History as an Interactive Digital Game. *The Australasian Conference on Interactive Entertainment*, Sydney, February 2004.

Ciger, J., et al., 2003. The Magic Wand. *The Spring Conference on Computer Graphics*, Budmerice, Slovakia, 2003, pp. 132-138.

CNR-ITABC, 2005. World Exposition on ancient Rome in VR: www.itabc.cnr.it/buildingvirtualrome/stampa.htm (accessed 21 June 2008)

Costabile, M. F., et al., 1998. Building Metaphors for Supporting User Interaction with Multimedia Database. *The 4th IFIP 2.6 Working Conference on Visual DataBase Systems*, L'Aquila, Italy, May, 1998, pp.27-29.

Crofts, N., et al., 2008. Current Official Version of the CIDOC CRM version 4.2.4 of the reference document. Definition of the CIDOC Conceptual Reference Model, http://cidoc.ics.forth.gr/official_release_cidoc.html (accessed June 2008).

Dix, A., et al., 2004. *Human-computer interaction (3rd ed.).* Harlow, England: Pearson Education Limited., 324 pages.

Fogli, D., et al., 2004. On electronic annotation and its implementation. *The Advanced visual interfaces*, Gallipoli, Italy, May 25-28, 2004, pp. 98-102.

Jablonka, P., et al., 2002. TroiaVR: A Virtual Reality Model of Troy and the Troad. *The CAA 2002 The Digital Heritage of Archaeology,* Heraklion, Crete, Greece, April 2002.

Lanier, J., 1992. The Top Eleven Reasons VR has not yet become commonplace: www.jaronlanier.com/topeleven.html (accessed 21 June 2008)

Lenzerini, M., 2002. Data integration: a theoretical perspective. *The ACM International Conference on Management of Data*, Madison, Wisconsin, June 3-6, 2002, pp. 233-246.

Leone, N., et al., 2005. Boosting information integration: The infomix system. *The Italian Symposium on Advanced Database Systems*, Brixen Bressanone, Italy, June 19-22, 2005, pp. 55-66.

Loscos, C. et al., 2004. The Museum of Pure Form: touching real statues in a virtual museum. *The VAST '04, 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*, Brussels, Belgium, December 7-10, 2004 .

Marcos, G., et al., 2005. Semantic based querying and retrieving information for artistic expressions: The Art-e-Fact Ontology. *The 4th International Workshop on Content-Based Multimedia Indexing*, Riga, Latvia, 21-23 June 2005.

May, W., 2005. Logic based XML data integration: a semi-materializing approach. *In Journal of Applied Logic*, Vol. 3, N. 2, 2005, pp. 271-307.

Mazzoleni, P. et al., 2004. CiVeDi: A Customized Virtual Environment for Database Interaction. *ACM SIGMOD RECORD*, Vol.33, N.3, September 2004.

Papagiannakis, G., et al., 2004. Real-Time Virtual Humans in AR Sites. *The CVMP 2004 1st European Conference on Visual Media Production*, London, UK, March 15-16, 2004, pp. 273-276.

Sinclair, P., et al., 2005. Concept browsing for multimedia retrieval in the SCULPTEUR project. *The 2nd European Semantic Web Conference*, Heraklion, Crete, Greece, May 29 - June 01, 2005.

Valtolina, S., et al., 2005. Dissemination of Cultural Heritage Content through Virtual Reality and Multimedia Techniques: a Case Study. *The 11th International MultiMedia Modelling Conference*, IEEE CS Press, Melbourne, Australia, January 12-14, 2005, pp. 214-221.

Valtolina, S., et al., 2007. Media for Knowledge Creation and Dissemination: Semantic Model and Narrations for a New

VSMM 2008
Digital Heritage – Proceedings of the 14th International Conference on Virtual Systems and Multimedia
M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis *Full Papers*

WEARABLE PERSONAL ASSISTANTS FOR THE MANAGEMENT OF HISTORICAL CENTRES

J. L. Izkara [a, *], X. Basogain [b], D. Borro [c]

[a] Building and Territorial Development Unit, LABEIN-Tecnalia, Parque Tecnológico de Bizkaia, Derio, Bizkaia - izkara@labein.es
[b] Escuela Superior de Ingeniería de Bilbao, EHU. Bilbao, Spain - xabier.basogain@ehu.es
[c] CEIT and Tecnun (University of Navarra). P. Manuel Lardizábal 15, 20018. San Sebastián, Spain, - dborro@ceit.es

KEY WORDS: Wearable Assistant, Multimedia Content, Web Services, Mobile Devices, Augmented Reality, Management of Cultural Heritage

ABSTRACT:

One of the main tasks facing the manager of a historical centre is to avoid degradation while retaining the historical value. For this reason, any intervention which takes place on the environment, should be carefully managed. Only when performing a proper diagnosis of the environment and its reality is possible to follow a high quality intervention. The integration of new information technologies has been crucial to the improvement of these processes providing new tools. Within the project called RASMAP, we have designed and implemented a mobile augmented reality platform based on a service oriented architecture. This project introduces the concept of Wearable Personal Assistant (WPA). WPA in the RASMAP platform represents an innovative wearable tool, which provides support to professionals in their daily activities (mechanical engineer, safety responsible person, diagnosis expert, etc.). This tool is based on augmented reality technologies, mobile devices and communication infrastructures. The development of the platform for the WPAs implies addressing several technological challenges: a) to overcome the limitations inherent in the mobile devices: speed, capacity of memory, capacity of storage, graphical features and others, b) to obtain tracking systems that they do not need to alter or to adapt the environment, c) to optimize for the transmission and reproduction of multimedia contents through wireless networks on mobile devices. In this article, we describe the RASMAP platform, as a basis for the development of WPA and the extension of its use for the management of historical centres. The quality and usefulness of the scientific-technological results provided by the WPA have been validated developing a demonstrator for the diagnosis of the conservation status of the historical centre of a small town in the Basque Country. The advantages to be gained by using WPA in the proposed scenario are among others: more efficient processes, improved communication between users, and local and distributed multimedia content records.

1. INTRODUCTION

Three technological factors can be identified in the favourable transformation of many of our daily activities, including process and habits of behaviour. These three factors are the reduced size and cost of the portable devices, the potential of the multimedia in mobile devices and the wide wireless connectivity of these devices. Today it does not seem strange to use mobile phones or PDAs to take a picture, send an e-mail or take notes in a meeting.

Mobile devices can become the key element to transform and improve various kinds of activities such as the visit to a museum, attend school or repair industrial equipment. They are being introduced in all areas of daily life. But the success of these applications lies not only in the technological evolution and their low cost, but also in the design and implementation of systems that support new mobile applications, which provide the user with support in his activities. Based on the evolution of technologies previously mentioned in this article, we introduce the concept of Wearable Personal Assistant (WPA) as a lightweight and versatile tool to support the daily activities of non-professional and professional users. WPA presented in this paper has been developed based on the mobile augmented reality platform implemented in the project RASMAP.

In the field of Cultural Heritage, one of the main tasks facing the manager of a historical centre is to avoid degradation while retaining the historical value. For this reason, any intervention which takes place on the environment should be carefully managed. Only when performing a proper diagnosis of the environment and its reality, is possible to follow a high quality intervention. The integration of new information technologies has been crucial to the improvement of these processes providing new tools.

The diagnostic processes require the participation of various experts from different fields of knowledge (architecture, history, environment, social sciences, etc.). This requirement makes it necessary to move all of them in the study area with a consequent cost. For the fieldwork user relies only on their knowledge, experience and ability to interpret the information it perceives. For this reason, the diagnostic procedure is very different depending on the expert who performs it. There is currently a shortage and a demand for standardization of these procedures. The concept of WPA represents a suitable solution for support to the users responsible for carrying out these tasks and standardizes its procedures. Besides it allows reducing the number of people who must move to place the diagnosis as well as their level of knowledge of different areas.

* Corresponding author.

2. RELATED WORK

Integrating mobile devices together with other information technologies can help to improve the effectiveness and convenience of information flow in various kinds of activities, for both professional and non-professional users. Thus, the mobile phone becomes the guide to the museum (Bruns, 2007) using its camera for the recognition of objects in combination with an extensive tracking system. The PDA (Wagner, 2006) could be the device to teach art history through an educational game based on collaborative augmented reality technologies Also the MIT Teacher Education Program (MIT, 2008) used the PDA technology in combination with GPS and Wi-Fi for the creation of simulation games that combine real-world experiences with additional information supplied to students by mobile devices.

2.1 Wearable computing and Mobile Augmented Reality

The concepts of wearable computing and augmented reality are well known and the system (Feiner, 1993) can be considered one of the first synergies of both technologies. The mobility requirements of users require abandoning systems based on desktop computer for mobile devices and then the first prototype applications of augmented reality-based mobile devices appear with Head Mounted Display (HMD) connected to a laptop, which carries out all processing, placed in a knapsack on their back (Höllerer, 1999; Piekarski, 2001; Piekarski, 2004). The following prototypes are starting to use PDAs (Wagner, 2003) or mobile phone (Möhring, 2004) as processing devices, and the displays devices are both video-through devices and see-through devices. Many technologies have been used to obtain a precise positioning application in augmented reality, all of them are complementary and there is no perfect solution (Izkara, 2007).

The projects that stand out in this area are MARS (Feiner, 1997), ARCHEOGUIDE (Project Archeoguide, 2008), Signposted (Wagner, 2003), client-server architecture for an application of augmented reality on PDA (Pasman, 2003), MR Virtuoso (Project Mr Virtuoso, 2008), Large Scale Mobile Phones Museum Guidance (Bruns, 2007) and ULTRA (Project Ultra, 2008) among others.

In (Weiser, 1991) Weiser defines the concept of "ubiquitous computing" as an environment in which computing technology is embedded into all kinds of everyday physical objects (such as appliances, doors, windows or desks). Recently miniaturized mobile devices have extended their capabilities from simple communication devices to wearable, networked computational platforms. Mobile AR can be viewed as the meeting point between AR and ubiquitous and wearable computing (Papagiannakis, 2008).

2.2 Mobile assistants for the management of historic centres

The idea to create mobile assistants is not new, it is running for many years and currently it has an important impulse due the potential benefits in many areas. The report (Freitas, 2003) resumes a comprehensive study about the development of wearable devices and personal data assistants where it can highlight the Xybernaut Mobile Assistant (Xybernaut, 2008) a commercial product widely available multi-purpose wearable device, educational applications such as 3COM Learning Assistant (3COM, 2008) and IBM web lectures delivering learning content to the learner through the mobile devices such as PDAs or mobile phones, or tools such as CyberTracker (CyberTracker, 2008) a software system developed for a PDA which enables trackers to record all the significant observations they make in the field, and the Electronic Guide (Electronic Guide Research Project, 2008), a tool for recording museum visits that use handheld devices to enrich learning experience for museum visitors. Also there are research projects oriented in this direction such as wearIT@work project (Project WearIT@atwork, 2008) which will develop a set of new solutions to support the workers of the future. WearIT@Work aims to develop a mobile computing platform that can empower professionals to higher levels of productivity by providing more seamless and effective forms of access to knowledge at the point of work, collaboration and communication, i.e. the aeronautic maintenance showcase (Giancarlo, 2006).

In the area of the management of historic urban centres the book (Pickard, 1997) examines key themes for the management of historic urban centres and also presents a variety of approaches utilised to face the different problems of the centres including management and regeneration action, environmental management and tourism and heritage management. A particular wearable solution for the cultural heritage sector (WearIT@CH, 2008) proposes a wrist wearable wireless computer addressing the needs of visitors will use the wearable device as a personalized multimedia guide, the needs of site guardians to support their routine on-site inspections and the needs of site managers and researchers will be able to gather all kind of data about visitors on-site behaviour.

Multimedia contents, interactive elements, audio-guides, etc. all of them are nowadays elements that have been incorporated, in a natural way, for the diffusion of cultural heritage and especially in museums. However the use of such technologies is poorly introduced in other activities in the world of cultural heritage and the potential of the technologies is quite high.

3. WEARABLE PERSONAL ASSISTANT (WPA)

3.1 What is a WPA?

We call Wearable Personal Assistant (WPA) a tool that provides user-support for the development of daily activities. A WPA has the following main features:

- Physically it is a lightweight device, easily manageable and with high quality visualization and audio playback devices.
- It provides support and assistance to professional users and non-professionals in their daily activities.
- It is a versatile tool, applicable to many different environments (education, industrial maintenance, cultural tourism, etc.).
- The type of information that provides a WPA is multimedia, quite visual and user friendly, with content specifically designed to assist the user in carrying out its tasks.
- The interaction with the tool is intuitive and limited in order to facilitate its handling by any type of user.
- WPA is an innovative tool that incorporates new technologies such as augmented reality, position tracking, image processing, wireless remote connectivity and multimedia content management services.
- WPA incorporates information personalized to the type of user (user profile, level of experience, etc.), contextualized to the activity being implemented as well as the user physical location (using positioning technologies).

3.2 RASMAP Platform

The implementation of the WPA is based on the mobile augmented reality platform developed in the RASMAP project (RASMAP, 2008). The RASMAP platform is based on a service-oriented architecture (SOA), where the final application is implemented from the services provided by different processes. For the definition of the required services we have analyzed comprehensively and systematically augmented reality (AR) scenarios in different sectors (cultural heritage, construction, maintenance and m-learning). For each scenario, first we identified the relevant group of services and functionalities required. On a second phase we established the integration of all identified services in the complete set of AR scenarios. These services were structured and organized in the following six large types: Interaction, Tracking, 3D Rendering, Multimedia presentation, Information management and Context-aware information.

A common interface or language has been defined for the request of the service and each process requesting this service will use it. In the case of the tracking service, for example, the information requested will be the position and orientation at a certain time. Tracking service will manage its resources and use the more appropriate technologies and devices to handle this request. The final application will receive the information about the position and orientation without being aware of the method used (WiFi-based, image processing, GPS, etc.) or even if it has been processed locally or remotely.

In the proposed approach, we present a mixed architecture, in which some of the detected services are executed directly in the mobile device, some others are remotely processed and some others will be processed locally or remotely depending on the context. The following figure shows the configuration of the architecture and services for the proposed platform.

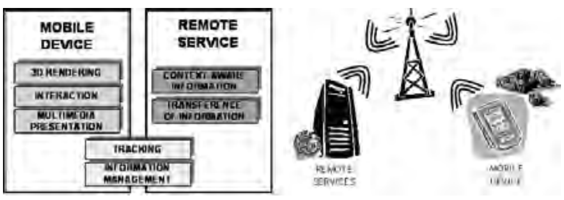


Figure 1: RASMAP platform

3.3 Software Components

The proposed design of the RASMAP platform has required developments in different research areas which include among others, graphics libraries for mobile devices, tracking libraries on 3D environments, image recognition algorithms based on 3D representations and multimedia transmission for mobile environments. The objectives proposed for these developments include the display of augmented reality applications on mobile devices such as PDAs, 3D positioning markerless on mobile devices, positioning the camera within the 3D environment, and the transmission of multimedia content.

The following sections describe various software components that perform several tasks including the augmented reality engine, robust positioning with occlusions in translations and rotations, and the creation and distribution of multimedia content in mobile environments. All the software components have designed for mobile devices based on Windows Mobile 5.0 and .NET Compact Framework.

3.3.1 Augmented Reality Rendering Engine: One of the main components of an augmented reality system apart from the image capture and the positioning/tracking is the 3D rendering. We have developed an engine for 3D visualization based on Vincent and PowerVR implementations of OpenGL ES; the first one is a software implementation and the second one is a hardware implementation.

Our selection is based on a low-lewel specification library such as OpenGL ES (OpenGL ES, 2008) (it is a subset of the OpenGL API for desktop environment), which describe the specifications for 2D and 3D graphics, and avoiding the high level rendering libraries available for mobile platforms due to the delay introduced in debugging and execution of applications on mobile devices and their incompatibility with the graphics standards. All the alternatives have in common that they use a scene graph to manage the 3D content.

We have developed a 3D visualization engine that incorporates interaction through buttons and touch screen, also displays complex 3D models generated using high-level 3D modelling tools and incorporates the text display using OpenGL ES techniques.

The image capturing module provides information to define the context to the system. In an augmented reality system real information is used to set the background of the augmented reality scene. For the implementation of this module we used both the Windows API and the SDK of the camera manufacturer. Figure 2 shows simple augmented reality scenes using the implemented 3D engine and the image capturing.

Figure 2: Augmented reality scenes

3.3.2 Tracking/Positioning: One of the major challenges of an augmented reality system is to achieve a good positioning of the user in the environment, providing both orientation and position with a high degree of accuracy (Kato, 1999). In order to make a robust 3D tracking, usually it is needed three main steps: a) camera calibration and initialization of the tracking system, b) initial camera pose (position and orientation) and c) tracking of the user movements relative to the initial or reference position.

For our developments, we have modified the ARToolkitPlus library (Wagner, 2007) for allowing partial occlusion of the markers, thereby improving the robustness of the system without adding markers (from now on, we will call the modified library as ARToolkitPlus'). Our implementation allows translation in the three axes (x, y, z) and rotation in z even when the marker is partially occluded. Apart from the processing algorithms mentioned, in order to soften the augmented model transformations we used a Kalman filter which avoid uncomfortable flickering for the rendering of the 3D scene. The following figure shows some screenshots of the achieved results.

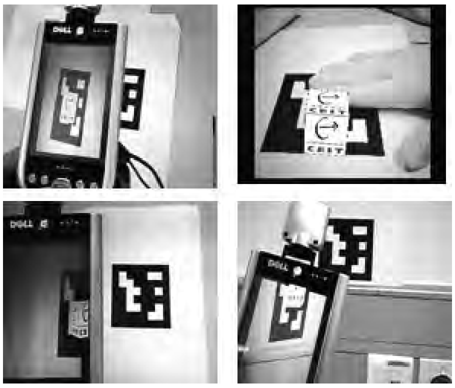


Figure 3: Using ARToolkitPlus'

3.3.3 Multimedia Management: This software component provides the user with a multimedia management service that lets her/him record video and audio, take pictures and draw or write over the photographs manually, and also write text notes. All these tasks are performed on a single screen with intuitive menus and organized in four tabs corresponding to text, audio, images and video management respectively. In addition, users can share their multimedia content or content from other users, sent or received via wireless (Wi-Fi or 3G) using the send/receive options enabled on all the menus of the multimedia management. The option send/receive is used in a mode transparent to the user to load and download multimedia content by web services located on a server that can be in the same building or thousands of miles away.

An additional functionality of the multimedia management component is the capacity to construct multimedia presentations based on the multiple data created or previously stored. This kind of service allows creating multimedia user's guides to assist the users in a rich media way. It uses an SMIL-2.0 player that presents all the multimedia contents (text, audio, pictures and video) in an integrated and synchronized form.

The advantages offered by this developed multimedia management with respect to other business applications such as windows mobile notes (text and audio), proprietary software camera (video and images), and utilities sending content via email are among others the following: a) integration into a single application. b) standard file formats (txt, wav, mp3, avi, mp4, wmv). c) multimedia content records: local storage and remote storage. d) collaborative content (send & receive). e) rich media presentation. Figure 4 illustrates different instants of taking a picture or making manual annotations, and the server hosting the web services of the collaborative mode.

Figure 4: Multimedia management system

The case study (Basogain, 2008) has been implemented in an university environment to demonstrate the suitability of this multimedia management in education, and has served as a testing ground for this software component. Any other scenario where the PDA becomes a real wearable personal assistant could have selected as a test ground, for example a visit to a museum or a cultural heritage site, a tour guide or a visit to an industrial exhibition, etc.

3.3.4 Remote Communication: The transfer of video and audio information in real time between two distant locations allows communication between two users who are in remote locations.

In our developments in the RASMAP platform we have integrated a video conference system for the remote communication. The system provides full duplex communication and it is based on either WiFi or 3G communication networks.

3.4 Hardware Devices

An augmented reality system consists of a group of devices with complementary functionalities connected and integrated through a software platform. From the hardware point of view the three main elements of the system are: The processing device, the visualization device and the positioning device. Alternatives for the three of them will be presented next.

The hardware device selected for the developments described in this article is the PDA Dell Axim x51v, mainly because its processing power, graphics card and screen resolution. For image capture, cameras Spectec SDC-001A and Spectec SDC-003A have been used. The only difference between both cameras is the resolution, 300KPixels and 1,3MPixeles respectively. The camera is attached to the PDA through the SD slot. The PDA has microphone and audio output to which are plugged conventional headphones to allow using the system in noisy environments. To give freedom the user to use his/her hands during the inspection, the PDA hangs on the user's neck through a band. The PDA screen itself, providing VGA resolution, is used as a display device. For the positioning, the system is based on marker-based optical tracking. Figure 5 shows a picture of the hardware components used in the described application.

Figure 5: Hardware devices

4. PROTOTYPE FOR THE DIAGNOSIS OF BUILT HERITAGE

4.1 Scenario Description

For the development of a first prototype of the WPA, we choose the task of diagnosing state of preservation and accessibility of buildings and public spaces of a historic centre, as one of the key tasks for the management of the historic centre of a village or city.

A user moves to the historic centre selected to make the diagnosis, in our case a small town in the north of Spain called

Labraza (see figure 6). The user is equipped with the WPA. The WPA includes a data sheet to be completed during the inspection.

Figure 6: User in Labraza

The task focuses on the diagnosis of three different elements of historic centre selected: A residential building that is attached to the wall (number 1034), a building of historic interest (El Portal Sur) and a public space (La Plaza Alta).

4.2 Methodology

Examples of the sequence of actions to be taken during the inspection are detailed below:

- The user arrives to the first element to inspect.
- Turns on the WPA, identifies himself on the system, selects the element and the system shows on the screen its corresponding diagnosis procedure.
- The system provides instructions prior to diagnosis such as how to place the marker, informs that it must take a photograph of the facade and presents the sheet to be completed to establish the diagnosis.
- During the diagnosis, the system provides additional information for each of the sections in the sheet:
    - Accessibility: Based on the marker position, when the user asks for additional information to diagnose the accessibility of the building, 3D information relative to the accessibility normative in buildings will be shown.
    - Vaults: An interactive 3D model library of the most common types of vaults will be shown.
    - Facade preservation status: In this case it comes to detecting if the facade suffers any kind of pathology. For this purpose there are presented to the user a photo and a brief textual and oral description with the most common pathologies.
    - Covers: With the objective to detect the kind of cover of the building, visual information of the most common types of covers is presented to the user.
- During the diagnostic process the user can download multimedia information about other inspections carried out previously. This information includes photographs and audio files. The user can make notes on the photographs for review in subsequent inspections.
- The user may request the assistance of a remote expert by using the video conference system included in the application.
- When the user has completed the inspection of the building he will have to move to the next item to inspect. The system will provide information on the route to the next inspection element with a 3D arrow accompanied by a three-dimensional reconstruction of the historic centre indicating the most relevant environment elements to help the user in his orientation.

4.3 Application Development

For the diagnostic procedure, we defined three main diagnostic elements: a residential building attached to the wall, a building of historical interest and a public area. On the other hand, we defined two different types of user: an expert user, and a non-expert user. For each of the elements and for each type of user a diagnosis procedure is defined. Table I shows an example of the procedure of diagnosis. In particular, the one corresponding to a non-expert user who is in front of a building of monumental interest:

ACTION	DESCRIPTION	TYPE
Place marker	Instructions to place the marker on the building to inspect	Multimedia Information
Initial instructions	Initial instructions for the diagnosis	Multimedia Information
Accessibility	Information about accessibility	Augmented Reality
Vaults	3D model library of different kinds of vaults	3D Model Library
Covers	Information about kinds of covers	Multimedia Library
Pathologies	Information about the facade preservation status	Multimedia Library
Route to next	Information about the route to the next inspection point	Augmented Reality

Table I. Procedure actions

Concerning the contents, several types of information have been used in order to define the contents shown to the user. The main functionalities provided by the system are described below:

- *Multimedia information reproduction*: This is a reproduction of a single media file, which consists of a picture, text and audio.
- *Multimedia Library reproduction*: It corresponds with the reproduction of a multimedia file consisting of a sequence of photos, text and audio associated with each picture.
- *Augmented Reality visualization*: This functionality displays the image captured by the camera in real time, augmented with a virtual 3D scene in a position and orientation relative to the position of the detected marker. This type of content is mainly used for routing the user between the inspection points and for assistance in the accessibility diagnosis. Figure 7 shows how the WPA assists the user in the diagnosis of the accessibility in the building entrance.

Figure 7: Augmented reality for accessibility diagnosis

- *3D Models Library visualization and interaction*: This system allows the visualization of 3D interactive models. The interaction implemented allows rotation of the model, as well as zoom in and out of the camera.

- *Remote Expert Assistance*: The WPA includes a video-conference system which allows the user making the diagnosis contact with a remote expert for assistance during the inspection. Figure 8 shows the user and expert during the remote assistance of the diagnosis.

Figure 8: Remote expert assistance

- *Multimedia recording, downloading and annotation*: The WPA provides the user with the multimedia management system. A repository of multimedia content allows the user access to the recorded actions undertaken during previous diagnoses, and also the user can add new multimedia records as it is illustrated in figure 9.

Figure 9: Multimedia files annotation

4.4 Test

The prototype has not been validated yet, however we have performed a first test with one external user. In this section we provide some information about the test done. This test allows us to detect some advantages and disadvantages of the developed prototype.

In order to perform the test we moved to Labraza. There were three people involved in the test session: The responsible person of the WPA, the WPA developer, and the user, who is an architect with long experience in diagnostics conservation status of historic buildings and environments.

Prior to the start the session, the user is informed about the tool and the objectives of the session, informing that he will be observed and his actions and comments will be recorded with a video camera.

The main advantages detected by the user were:

- The manageability of devices exceeds expectations, as it can hang the PDA on the neck that gave freedom of hands movement to the user.
- The interaction with the application was simple and intuitive.
- Information provided was adequate and very interesting.
- The information presented using 3D models provided enough help and was of great interest for the user.
- The inclusion of the marker corresponding to the diagnosis element in the diagnosis sheet was very interesting and useful.
- Information given using 3D models with the aim of guiding the user in the environment was really useful.

The contents of augmented reality shown are very useful in the tasks of accessibility diagnosis and guidance to the next element. In the first case, the system allows the user to visually access the accessibility rules and make a diagnosis without measure. For guidance to the next point, visual information provided allows the user to orient easily within an unknown environment, as well as to reach the next inspection point in a simple way.

The tool for managing multimedia content provides apart from a mean of presentation of rich multimedia content that help during the diagnosis, other features like taking pictures, annotations over them, recovery of photographs taken during previous inspections, upload photographs for upcoming inspections and so on. These actions are also possible with other types of multimedia content such as video, text or audio.

Remote communication with an expert facilitates the diagnosis by non-expert users since the system allows accessing experts during the performance of the task.

However, problems were detected; for example:

- Light intensity on the screen of the PDA under conditions of high brightness outdoors is a bit poor.
- Battery life of the PDA during the assessment session was about 1 hour and 45 minutes, which is not enough to make the diagnosis of a complete historic centre.
- The user requests for the possibility to have the diagnosis sheet integrated in the WPA, which is not yet implemented.

5. CONCLUSIONS AND FUTURE WORK

In this paper we present the concept of Wearable Personal Assistant as an innovative tool for the assistance to professional and non-professional users in their daily activities. The concept of WPA has been defined in the context of a project called RASMAP in which we have developed a mobile augmented reality platform based on a service-oriented architecture.

Several software components have been developed for the implementation of the RASMAP platform. We have implemented an augmented reality rendering engine for mobile devices based on software and hardware implementations of OpenGL ES. Based on ARToolKitPlus, we implemented several algorithms to improve the robustness of such marker based tracking library using mobile devices. We have implemented also a multimedia management tool; this tool it is based on a set of web services and provides functionalities to record, edit, download and upload multimedia contents. In the RASMAP platform we have also integrated existing tools such as a video conferencing system for the remote communication between users remotely located.

The RASMAP platform is the technological support for the development of WPA. We have developed a prototype for the diagnosis of the built heritage. The tool has been tested on site by an architect with long experience in the diagnosis activities. Regarding the usability of the tool, although several problems were detected, the user perceived an added-value in the use of such tool for their daily activities. Among others, the standardization in the diagnosis processes, the reduction of time and the number of experts needed to be moved to the site to make the diagnosis are some of the most important benefits of such tool.

The use of tools such as WPA based on new technologies is not only a instrument for the diagnosis of the conservation status of the built heritage, but also can be very useful for many other tasks in the management of cultural heritage sites, for example: as a supporting tool for the visualization and selection of alternatives in the revitalization processes of the historical sites, also in the phase of maintenance of building or infrastructures etc. For non-professionals the concept of WPA could be applied to aid the user in the visit to a museum or other cultural places. As previously mentioned cultural heritage is not the only sector to apply such novel technologies. We are starting to develop a new prototype for professionals in the areas of mechanical maintenance.

In the near future we plan to extend the services of the RASMAP platform, developing new technologies for improving the tracking. We will develop new algorithms for the vision based tracking and combine it with other different technologies (GPS, WiFi, etc.). We plan also to development of new services for context awareness, considering both contexts: the user and the environment. These services will allow the system to provide personalized information to the user.

For the prototype in the diagnosis of the built heritage we will include minor changes in the application according to the feedback provided by the user during the test and we will validate the tool using the developed WPA for the diagnosis of a historical centre comparing the results with the traditional methods.

REFERENCES

Basogain, X. 2008. Olabe, M., Espinosa, K., Rouèche, C. and Olabe, J.C. Collaborative Multimedia Mobile Tool for Teaching Laboratories. *8th International Conference Developing Innovative Video Resources for Students Everywhere*, Diverse' 2008. July 2008, Haarlem, INHOLLAND University, The Netherlands. (Accepted)

Bruns, E. 2007. Brombach, B., Zeidler, T. and Bimber, O. Enabling Mobile Phones To Support Large-Scale Museum Guidance. In *IEEE Multimedia*, vol. 14, no. 2, pp. 16-25.

CyberTracker 2008. www.cybertracker.org (accessed 16 June 2008)

Electronic Guidebook Reseach Project 2008. http://www.exploratorium.edu/guidebook/ (accessed 16 June 2008)

Feiner, S., 1993. Macintyre, B., Seligmann, D. Knowledge-based augmented reality. *Communications of the ACM archive.*Volume 36, Issue 7, Special issue on computer augmented environments: back to the real world. pp 53-62.

Feiner, S. 1997. B. MacIntyre, T. Höllerer, and T. Webster. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Proc. ISWC '97, October 13-14*, Cambridge, MA.

Freitas, S. 2003. and Levene, Mark. Evaluating the development of wearable devices, personal data assistants and the use of other mobile devices in further and higher education institutions. JISC Technology and Standards Watch Report (TSW0305), pp. 1-21.

Giancarlo, B. 2006. Andrea Lorenzon, Nicolas Chevassus, Valerie Blondel. Wearable Computing and Mobile Workers: The Aeronautic Maintenance Showcase in the WearIT@Work Project. *Third International Forum on Applied Wearable Computing* IFAWC2006 March 15-16, Mobile Research Center, TZI Universität Bremen, Germany.

Höllerer, T. 1999. Feiner, S.; Terauchi, T.; Rashid, G.; and Hallaway, D.: Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System. Computers and Graphics, 23(6), pp. 779-785.

Izkara, J.L. 2007. Pérez, J.; Basogain, X.; Borro. Mobile Augmented Reality, an Advanced Tool for the Construction Sector. *D. Bringing ICT knowledge to work: Proceedings of CIB 24th W78 Conference* in Maribor 2007, pp. 453-460.

Kato, H. 1999. and Billinghurst, M.: Marker Tracking and HMD Calibration for a Video Based Augmented Reality Conferencing System. Proceeding of the 2nd Internation Workshop on Augmented Reality, pp 85-93.

MIT 2008 Handheld Augmented Reality Simulations. http://education.mit.edu/ar (accessed 16 June 2008)

Möhring, M. 2004. ; Lessig, C.; and Bimber O. Optical Tracking and Video See-through AR on Consumer Cell-Phones. In: *International Symposium in Mixed and Augmented Reality*, Arlington, USA, pp. 252-253.

OpenGL ES 2008. http://www.khronos.org/opengles (accessed 16 June 2008)

Papagiannakis, G. 2008. Singh, G.; Magnenat-Thalmann, N.; A survey of mobile and wireless technologies for augmented reality systems; Computer Animation and Virtual Worlds, Volume 19, pp 3-22.

Pasman, W., 2003.. and Woodward, C.: Implementation of an augmented reality system on a PDA. Poster article in *Proceedings of International Symposium of Mixed and Augmented Reality*. ISMAR 2003, Tokio, Japan.

Pickard, R. 1997. *Management of Historic Centres*. Edited by Robert Pickard ISBN: 978-0-419-23290-2

Piekarski, W. 2001. and Thomas, B.: Tinmith evo5 – An Architecture for Supporting Mobile Augmented Reality Environments. *Proceedings of the International Symposium on Augmented Reality* ISAR01. New York, USA.

Piekarski, W. 2004. and Smith, R.; and Thomas, B. H.: Designing Backpacks for High Fidelity Mobile Outdoor Augmented Reality. *Proceedings of International Symposium on Mixed and Augmented Reality* ISMAR04, Arlington.USA.

Project ARCHEOGUIDE 2008. http://archeoguide.intranet.gr (accessed 16 June 2008)

Project Mr Virtuoso 2008. http://studierstube.icg.tu-graz.ac.at/handheld_ar/virtuoso.php (accessed 16 June 2008)

Project ULTRA 2008. ULTRA Light-weight Augmented Reality Mobile System. http://ist-ultra.org/publications.shtml (accessed 16 June 2008)

Project WearIT@work 2008. http://www.wearitatwork.com/ (accessed 16 June 2008)

RASMAP 2008. RASMAP Project. http://www.labein.es/rasmap (accessed 16 June 2008)

Wagner, D. 2003. and Schmalstieg, D.: First Steps towards handheld augmented reality. *Proceedings of International Symposium of Wearable Computers* ISWC03, New York, USA.

Wagner, D. 2006. Handheld AR for Collaborative Edutainment. Lecture Notes in Computer Science, Springer Berlin / Heidelberg. ISBN 978-3-540-4977-9. pp. 85-96.

Wagner, D. 2007. and Schmalstieg, D.: ARToolkitPlus for Pose Tracking on Mobile Devices. *Proceedings of 12th Computer Vision Winter Workshop (CVWW'07).*

WearIT@CH 2008 Wearable Solutions for the Cultural Heritage Sector. http://www.wearitatwork.com/wearIT-CH.259.0.html?&0= (accessed 16 June 2008)

Weiser, M. 1991. The computer for the 21st century. Scientific American, 265(3). pp. 94-104.

Xybernaut Mobile Assistant 2008. http://www.xybernaut.com/ (accessed 16 June 2008)

3COM Learning Assistant 2008. The 3Com University http://www.3com.com/support/en_US/training/ (accessed 16 June 2008)

ACKNOWLEDGEMENTS

The work described in this paper is based on a project partially funded by the Spanish Ministry of Education and Science. The project is called RASMAP - TIN2006-15418-C03 and participant institutions are LABEIN-Tecnalia, CEIT and Tecnun of the University of Navarra, and Engineering School of Bilbao UPV-EHU.

AUTHOR INDEX