The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Haptic Navigation and Exploration of High Quality Pre-rendered Environments

R.G. Laycock[†,1] S.D. Laycock[1] and A.M. Day[1]

[1]School of Computing Sciences, University of East Anglia,
Norwich, NR4 7TJ, UK

Abstract

Visualising in real-time high quality virtual environments, which are suitable for cultural heritage and virtual tourism, is often a challenging endeavour. This is primarily, due to the cost of rendering complex architectural structures. Incorporating the sense of touch into real time environments enhances a user's experience, leading towards a higher level of immersion. However, the confining requirement of the haptic feedback loop to cycle at 1000Hz has led to many applications employing lower quality scenes. In this paper a technique is presented which permits high quality pre-rendered animations of dynamic environments to be both visualised and navigated at high interactive rates. In particular, the approach provides the user with the ability to touch the content of the animations and to freely orientate themselves in any direction, leading to haptically aware movies. Consequently, this permits an exploration of an otherwise prohibitively complex scene consisting of large volumes of geometry and texture maps, combined with realistic lighting models. This component is illustrated within a virtual tour framework enabling users to navigate semi-prescribed routes using haptic feedback both in a spatial and temporal context.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Interaction Techniques

1. Introduction

Virtual environments provide an excellent domain in which to visualise archaeological artefacts. Unlike the traditional museum environment, where an artefact is frequently displayed out of context in a glass cabinet, the virtual environment enables the object to be digitally reconstructed and reunited with additional objects that may be stored at different geographical locations. Furthermore these objects may be visualised in the correct context and augmented with additional media to enhance the user's experience. Besides the visual impact obtained from displaying a set of objects within a geographically and temporally accurate virtual environment, additional information may be perceived via the incorporation of a haptic feedback device. A haptic feedback device provides the user with the ability to touch three-dimensional virtual objects and to feel their shape and contours. This is particularly important in obtaining a better understanding of an object's three-dimensional form.

To maintain a smooth and continuous interaction between the haptic feedback device and the virtual objects the feedback loop is required to cycle at 1000Hz. Consequently the graphical rendering is typically compromised, reducing the scene's fidelity to permit the desired interactive rate to be achieved. In this paper an animation is adopted to represent the virtual environment. It thus permits large three dimensional environments, consisting of millions of vertices and many textures, to be rendered offline under computationally expensive illumination effects. By employing animation in this way the user is provided with high quality imagery representing the objects within a virtual environment. To gain further insight into an object's structure techniques are developed to enable the user to touch the objects displayed within the environment via a haptic feedback device. By giving the user the ability to touch the virtual environment an increased level of interactivity will be realised.

† rgl@cmp.uea.ac.uk

The main contribution of this paper is a novel technique that enables users to interactively navigate and explore a high quality pre-rendered environment with haptic feedback. This is particularly pertinent for virtual tours of cultural heritage locations, where obtaining high fidelity images at the desired interactive frame rates is difficult to achieve.

2. Previous work

There are several published works on the concept of representing a virtual museum, [BBF*02] [OBP05] [Bre01]. In particular Bergamasco et al. in [BBF*02] describe the Pure Form project, which is motivated by the requirement to include haptic feedback and therefore enable the full appreciation of three-dimensional artefacts. McLaughlin et al., in [Bre01], agrees with this claim adding that by creating programs with haptic feedback a user is able to touch fragile objects without adversely effecting their conservation. However, these approaches require a computer generated image to be rendered at interactive frame rates. Even with the progression of advanced graphics hardware technology simpler virtual environments are inevitable. These comprise of smaller texture maps and models lit with lower quality lighting calculations, [LWR*03] and this leads to a restriction when attempting to interact with large high fidelity scenes in real time.

In [MJ06], Morris et al., targeted this issue by combining a single ray-traced image with haptic feedback for an abstract scene. Their technique was driven by the concept that in a haptic feedback environment the viewpoint rarely alters, suggesting that views from static locations from within a scene are acceptable. However, in a virtual tour of a cultural scene, it is a requirement that the user is able to traverse the scene and stop to touch interesting artefacts along the way. For instance, it would be beneficial to observe a particularly ornate sculpture, in context, attached to the façade of a building. During the traversal of the environment further insight into the three dimensional work undertaken into the decorative sculpture could be gained through haptic interaction.

The functionality provided for the programmer to record the haptic device's movements has been utilised by the haptics community for many applications including virtual training of surgeons [WSJH04] and rehabilitation for stroke patients [RMJ*05]. The ability to record the device's movements and to play them back has been exploited by [GMS06] for enabling users to receive pre-recorded haptic feedback whilst watching a movie. The haptic feedback paths were manually constructed to match an action in the movie. The system did not allow direct contact between the device and the content of the movie.

3. An approach to exploring and navigating a pre-rendered scene

This section provides details on exactly how the methods have been implemented. Initially a computer graphics (CG) representation of the area to be modelled must be created. The CG representation may then be used to render movies with different paths throughout a scene. The focus of the method is to allow the user to interact with some of the objects displayed in the movie utilising a haptic feedback device. The next three sections in this paper discuss how a navigable movie can be created, how the touchable content is stored and how the haptics may be incorporated. In addition to touching the objects within the movie the user may use the haptic device to orientate themselves and to also modify the speed at which the movie is playing. It has been shown that adding these physical methods of interaction assists the user with media control [SMS*01]. Snibbe et al. [SMS*01] implied that since humans are used to interacting with visual media via physical systems techniques should be developed which replaces the physical sensations that are lost when interacting with the generic keyboard and mouse. In this section techniques for using the Phantom Omni device for controlly head orientation and the movie play speed are discussed. In addition to this the haptic device is utilised to manipulate a user interface designed specifically for cultural heritage applications.

3.1. Navigable movie creation

Computer generated scenes are advantageous for cultural heritage since they provide the ability to reconstruct environments that no longer exist. In this paper three-dimensional virtual environments are constructed from archive data using 3ds Max. The interactive modelling package is sufficiently powerful to enable the construction of large environments. Furthermore, it has the capability of importing a wide range of common model formats that may be used to store the large volume of digitised artefacts that exist today. To provide the context in which the cultural heritage site may be visualised an automatic urban modelling technique is adopted, [LD03]. This approach utilises building footprints and road layout information in conjunction with height data to automatically construct the terrain and the buildings. Each building's cross section is specified by its footprint and it is converted into a three dimensional volume by incorporating the height data. Various roof modelling techniques are employed to provide each building volume with a roof model based on its cross section and height. The resulting geometric model is subsequently improved via the integration of ground level images, which were previously compiled into a texture library.

A significant factor contributing to the fidelity of a virtual environment is the illumination that is directed upon it. Many different light sources are required ranging from the light induced by the sun and sky for exteriors to interior spot lighting. Many surfaces are required to accurately represent the interactions between these light sources and the objects in the three dimensional models. Consequently the geometry of the scene increases. A direct result of increasing the light sources and polygon count of the scenes is an increase in rendering time. Therefore the proposed solution is to maintain the visual quality obtained using an offline renderer and display the results in real time using an image based rendering technique.

Chen's paper on QuickTimeVR, [Che95], describes a number of techniques that enable the visualisation of offline rendered movies in real time. A particularly interesting theoretical concept presented involves the capture of many cube maps at small intervals along a fixed path. The cube maps are consequently switched in and out to enable the user to perceive a continuous motion. This approach is well suited to the application of a virtual tour since it enables a user to follow a fixed guided tour path leaving them to concentrate on interacting with their surroundings.

Using 3ds Max a spline is constructed and six cameras are attached to follow it. Each of the six cameras has a field of view equal to 90 degrees and is orientated to point along the six axes of a three dimensional coordinate system. To promote the easy generation of the six respective animations MaxScript is written to enable the automatic construction of the six cameras and the rendering of the corresponding movies.

The real time rendering component is implemented using OpenGL. It consists of a cube, with each face being textured with a frame from the corresponding movie. During rendering the camera is positioned in the centre of the cube and as time passes the appropriate frames from the movie are extracted and applied to the faces. In accordance with Chen's theoretical concept this simulates the effect of a virtual tour on a vehicle, where a user travels along a predefined path but is free to obtain a new view from any orientation. As the user changes their orientation view frustum culling is employed to ensure only the visible faces of the cube are updated. This reduces the time required to decode the necessary frames from the movie files.

3.2. Creating touchable content

By incorporating a force feedback device the user is able to physically interact with the virtual world. It enables the user to explore the scene and appreciate the three dimensional structures contained within the movie. This section discusses how the haptic representation is constructed leading to *haptically aware movies*.

A fundamental component of the haptic feedback system is the knowledge concerning the location and topology of the touchable objects. Therefore it is desirable to obtain information from 3ds Max concerning the position of both the camera used to acquire the movie and the touchable objects. This information is compiled using MaxScript to acquire the six degrees of freedom translation and rotation parameters over the entire timeline of the animation.

First the virtual camera is selected and the MaxScript is run to convert the coordinates from the right handed coordinate system of 3ds Max into the left handed coordinate system of OpenGL. The converted coordinates are exported into a file for use by the real time haptic rendering component. This procedure is repeated for each touchable object in the scene. In addition each touchable object is exported in .3ds format to enable the structure of the objects to be known. The objects which will be touched are not required to be the same as the objects which are visualised. Section 3.3.1 provides more information on the permissable representations for the touchable objects.

The haptic rendering system imports the three dimensional objects along with the camera and object transformation parameters. These are used in the following way to determine how each object is related to the camera. A transformation matrix, C_{tm}, is stored for the camera in world coordinates. A transformation matrix, O_{tm}, for every object is also stored in world coordinates. Figure 1(a) illustrates these matrices in respect of one object and the camera. The origin of the scene used for the visualisation and the haptics requires the viewer to be positioned at point(0,0,0). For this reason the camera is set at point(0,0,0) and the objects in the scene are transformed relative to the position of the camera. To acheive this an object can be transformed by $local_{tm}$, which is calculated as follows. Figure 1(b) depicts the object in relation to the camera.

$$local_{tm} = C_{tm}^{-1} O_{tm} \tag{1}$$

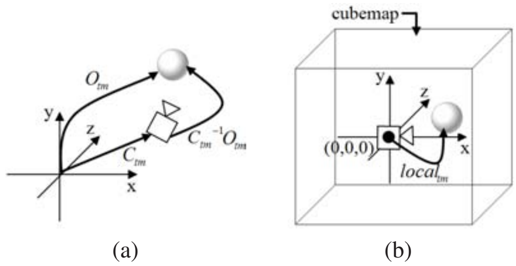


Figure 1: *(a) depicts an object and a camera in world coordinates. The matrices that can be used to transform them are illustrated. (b) illustrates the same objects positioned relative to each other in a coordinate frame that is local to the camera. The cubemap, which will display the six images from the movies, is depicted in figure (b) positioned at point (0,0,0).*

This is only part of the solution since the user is permitted to rotate their head as they navigate. The angles of rotation must be taken into account. The transformation matrix for the orientation of the head can be multiplied by $local_{tm}$

to obtain a new matrix, $ObjLocal_{tm}$, which can be used to transform all the objects in the scene. Once the objects have been transformed relative to the camera, methods can be employed to determine appropriate forces resulting from user interactions. The next section discusses how the haptic feedback is incorporated.

3.3. Incorporating haptic feedback

Exploiting the sense of touch in virtual reality simulations strives to provide a more immersive user experience and has been shown to improve interaction tasks [AKB01] [Che99] [CBG00]. In this paper haptic feedback is introduced into two areas; primarily for the navigation and exploration of the three dimensional scene and secondly to interface with the maps and the 2.5D user interface. This section focuses on the primary with the secondary being discussed in Section 3.6.

3.3.1. Touching the virtual objects

For many years users have been able to touch polygonal objects using a single point [ZS95] [RKK97]. This is akin to touching the objects with the tip of one finger. However, this is not particularly natural, since users often require touching an environment with an object of finite size, such as a tool. To achieve this a method similar to the Voxmap-PointShell approach may be exploited [MPT99]. Enabling a tool to be used in the environment significantly enhances the variety of applications. The three dimensional object manipulated by the haptic feedback device will be referred to as the haptic tool. The process of determining the appropriate forces for a given position of the haptic tool is known as haptic rendering. A typical haptic rendering approach requires the following stages: collision detection, contact determination and restoring force calculation.

All the touchable virtual objects incorporated into the scene are constructed from triangles and are typically exported from 3ds Max, as discussed in Section 3.2. A test point approach has been developed for the haptic rendering. A simplified version of the haptic tool is created where the vertices are strategically placed. The location of the test points can be determined either by using a spatial partitioning approach or by obtaining vertices on the convex hull. To improve the efficiency of the collision detection each object in the scene is divided up into a three dimensional grid of cells in a precomputation phase. A list of triangles that pass through each cell are recorded with each cell. This information can be stored in an array, Spatial_Array. To perform the collision detection between a test point, *tp*, and a virtual object several queries are employed. The first query is devised to trivially reject a possible intersection by determining if the test point is inside the extents of the virtual object. Since the calculations for collision can be performed local to the virtual object its extents can be calculated once. The variables

```
Begin Procedure: Test Point Collision Detection
    If test point, tp, is outside the extents of the Object THEN
        return NO_COLLISION
    x := (tp_x - ObjectMinX)/ObjectCell_Width
    y := (tp_y - ObjectMinY)/ObjectCell_Height
    z := (tp_z - ObjectMinZ)/ObjectCell_Depth
    index := x*NoCells*NoCells + y*NoCells + z
    If Spatial_Array[index] ≠ NULL THEN
        Determine the closest distances from the
        tp to each triangle in the cell.
        If closest distance < ε THEN
            return COLLISION & SCP
    return NO_COLLISION
End Procedure:Test Point Collision Detection
```

Figure 2: *Collision Detection procedure for a test point colliding with a polygonal object*

ObjectMinX, *ObjectMinY* and *ObjectMinZ* store the coordinates of the minimum extent. Similar variables are required for the maximum coordinate. If the test point is inside the extents of the object the cell in the Spatial_Array that encloses the test point can be calculated. These stages can be computed efficiently, using simple formulae, and are independent of the number of triangles in the object. To obtain the point of collision between the test point and the object the closest points are determined between the test point and each triangle stored in the selected object cell. A surface contact point, SCP, is returned if the closest point to the test point is within a small threshold. The complete procedure is described formally in Figure 2. Once the collision has been detected and the surface contact point and features obtained, a constraint based haptic rendering algorithm [HBS97] is be performed on a test point basis. The resultant force is obtained by adding all the forces acting on each test point in collision. To obtain high quality touchable objects a graphical representation of the virtual objects imported from 3ds files will not be illustrated to the user. Instead the virtual objects will be transformed so that they appear at the same location as they are in the movie. This is acheived by multiplying the geometry by the transformation matrix, $ObjLocal_{tm}$. The way in which this matrix is constructed was discussed in Section 3.2. This matrix will then be used to transform the test points of the haptic tool into the local coordinate frame of each virtual object in turn. By keeping the collision detection test in the virtual object's local coordinate frame the precomputation phase which created the grid cells can be utilised.

Since the virtual objects are not rendered graphically the user can touch the objects illustrated in the movie and thus touch high quality rendered objects. Additionally, this has the advantage of being able to use a different geometric representation for the touchable objects compared to those that were utilised in 3ds Max at the point of rendering the movies. However, the geometric representation's silhouette must be preserved in comparison to the silhouette of the objects used in the movie. The next section will make the reasons for this clear. The separation between touchable geometry and graphical geometry provides flexibility which can be exploited for haptically rendering complex virtual objects. Furthermore, by not rendering the touchable objects the high quality pre-rendered movie is not compromised by the appearance of lower quality objects rendered in standard OpenGL. The next section discusses some additional factors that are taken into account to strive for seamless integration of the haptic tool within the pre-rendered environment.

3.3.2. Integrating the haptic tool

In this section techniques to integrate the OpenGL rendered graphical representation of the haptic tool with the pre-rendered movie will be presented. If the representation of the haptic tool looks out of place when visualised against the movie then the sense of immersion will be lost and a significant aim of the technique will be hampered. Two areas for seamless integration will be investigated namely, lighting and occlusion. These are motivated from Morris et al. [MJ06] where the haptic cursor was integrated into a single ray-traced point-based rendered image. To obtain a close match to the illumination, the lights used in the 3ds max scene can be exported with the cameras. Although the effects of these lights within a 3ds Max rendered scene are calculated differently to OpenGL lighting at least the positions can be obtained to ensure the specular highlights appear appropriate on the graphical rendering of the haptic tool. More importantly, is the ability to enable the haptic tool to be occluded by objects in the movie, since the haptic tool must be able to disappear behind the touchable objects within the scene.

During a prescribed route a selection of objects move in and out of the area that can be reached by the haptic tool. It is only necessary to consider objects for occlusion which are within range of the haptic tool. Section 3.4 describes how the list of touchable objects can be obtained. Although the geometric representation of the object will not be displayed it can be rendered into the depth buffer by disabling the OpenGL color buffer for writing when rendering the virtual objects. The haptic tool is subsequently rendered and the hardware accelerated z-buffer algorithm will deal with the occlusion problem simply and efficiently. This method relies on the silhouette of the geometric representation of the virtual objects matching sufficiently to the representation used in the generation of the movie. If the silhouettes do not sufficiently match then the tool will be occluded by the virtual objects incorrectly.

The haptic interaction can be utilised for other functions beyond simply touching the objects within the environment. Although touching them is particularly important when obtaining information regarding weight, surface contours, size and three-dimensional information. Section 3.5 provides additional functions of the haptic device within the movie context.

3.4. Determining touchable objects

Given a particular path through the virtual environment only a subset of the objects visible in the scene will actually be candidates for haptic interaction. Consequently for each frame of the prescribed route a list is built that contains an index into an array of touchable objects. For each frame all the object's transformations are obtained and combined with the camera's transformation to determine the objects position relative to the camera. This information is utilised to permit a sphere check to determine if the object is touchable by the haptic tool at a given frame. This results in a list of touchable object indices for each frame of the movie. To reduce the memory requirement frame coherence is exploited allowing the list to be converted into a table of touchable object sets and a lookup table. This is achieved by reducing the original collection of touchable object lists into a unique set. A lookup table is constructed identical in size to the number of frames of the movie and each location of the table refers to one of the unique touchable object lists.

By determining the touchable objects as a preprocessing step the location of the touchable objects is known for each frame during haptic rendering in constant time.

3.5. Utilising haptic feedback for navigation

One challenge regarding the presented technique is to provide the user with a sense of immersion despite the fact that a section of pre-rendered movie is being played to them. Touching the virtual objects is a significant factor in providing this sense of immersion, however, the freedom to control their movement is also important for an immersive interactive application. The direction of travel within a pre-rendered movie can not be modified but the view direction and the speed of movement is permitted to be altered. In Section 3.6 a technique will be presented to enable the user to take control of the path they take. To avoid the use of additional input devices the view direction and playspeed will both be modified using the haptic feedback device. When the user presses and holds the white button on the device a virtual spring is initialised with one of its end points clamped at the current cursor position. As the user moves from this point the spring extends and once its length is greater than a certain threshold the scene will rotate. Two springs are utilised one for the x component and one for the y component, which perform scene rotations in the x and y axes respectively. A force is calculated based on the extension of the springs. The user is then able to link the magnitude of the force with the speed of rotation. The play speed of the movie can be modified in a similar fashion except that the blue button can be held. When either of the device's buttons are depressed the virtual objects are no longer touchable. This is important to prevent confusion over forces generated from the interface functions and the interaction forces within the environment. The visual feedback utilised to exhibit the change of haptic function involves altering the haptic tool from a hand when touching virtual objects to 3D cursors for movie controls.

3.6. A haptic enabled user interface for a virtual cultural heritage tour

The proposed virtual cultural heritage tour enables a user to navigate a large area visualising artefacts within their correct geographical location. The nature of a guided tour system is exploited to enable high quality cube maps to be pre-rendered for each frame of the movie. This concept permits a user to look in any direction as they travel along their chosen route. Integrating the branching movie framework, [Che95], permits a user to navigate between the different prescribed paths through the environment. The branching movie framework describes the early computer graphics technique employed by the games industry, comprising a directed graph where edges represent movies and vertices represent connections between the movies. In the proposed technique many movies are generated and unlike the traditional branching movie framework connections are permitted at any point, owing to the nature of the cube map movie frames. Therefore users are able to select their choice of direction at key junction points in a seamless manner. For example, in a street scene key points would be located at junctions in the road network as well as at the entrance of buildings such as St Andrew's Hall.

Haptic feedback is incorporated to enable the routes through the scene to be selected, the velocity of the camera to be altered and the objects in the scene to be explored. By utilising the haptic feedback device for the entire user input the application is suitable for a standalone installation in a room with low ambient lighting. Exhibiting the work in this environment is particularly well suited to a kiosk in a museum and enhances the visual quality perceived on a standard monitor. Furthermore since the keyboard and mouse are no longer required, low lighting is acceptable. In addition the development of cheap consumer level 3DOF output haptic feedback devices ensures that a system consisting of a standard PC, haptic device and monitor is economically viable for both museums and education facilities.

During the running of the system the user is presented with further information concerning their current position within the cultural heritage site, as well as provided with context information with respect to the touchable objects. Both the temporal and spatial context information is in the form of a variety of media including short animations, images and text. The remaining two subsections provide details of the data structures required to maintain data concurrency and integrity, followed by the user interface design and implementation specifics concerning the inclusion of haptic feedback.

3.6.1. Data structure

Figure 3 presents a class diagram illustrating the overview of the data structure developed. At the root of the data structure is the location class. A location contains a map of the area it covers plus a list of sites. Each site refers to a particular position on the location map and may have associated with it a name, movies and images that can be viewed from that position. A site may share this media with other sites at the same location. Therefore the location class contains a unique list of movies, objects and images and the list of sites indexes into it.

The MovieSpline class stores an index into the list of cube map movies. Each movie enables the user to seemingly travel from a particular location on the map to another location. A set of control points are stored and displayed on the map using a Catmull-Rom spline. This permits a marker to travel along the spline illustrating to the user their current location. The user can select these points to enter into the movie at their locations. A list of integers indexing into the content and 3d object lists is maintained. Associated with each index is a valid frame range to ensure that 3D objects and image based content are only updated within the valid frame interval. This provides context information to the user as they progress along the timeline. Furthermore, each object has an index associated with it to allow objects to display context information when inspected by the haptic tool.

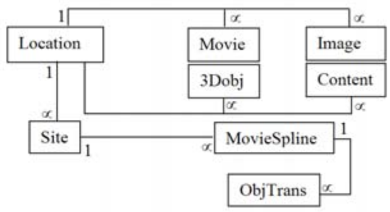


Figure 3: *A class diagram for the data structures of the system.*

3.6.2. Haptically enabled user interface

The user interface is divided into three main areas: the map area, content area and movie area. It is manipulated in two modes using the haptic feedback device. The most important mode is to manipulate the haptic device within a three dimensional volume for the movie exploration. However, the device is also used to select buttons and control the user interface in 2.5D. Initially the user manipulates a 2D cursor which is constrained to a plane giving the illusion of a flat area covering the entire screen. Areas of the interface are divided with small walls which keeps the user in particular areas as they slide over the interface. Each map or content area may be panned horizontally and vertically by pressing the blue button down and dragging the device. A force is perceived as the cursor moves. The content in these areas may be scaled by holding the white button down and pushing the device into the screen. A key component of the interface is to enable the transition of the haptic device from the 2.5D interface mode to the 3D movie mode. This is acheived by pushing the device into the screen when the 2D cursor is displayed on top of the movie. The user must pull the device towards them to return to the 2.5D mode from the 3D movie mode.

Figure 4 shows the interface of the program with the hand cursor currently interacting with the objects displayed within the movie. The movie depicts a highly detailed model of the interior of St Andrew's Hall. Four objects were inserted into the model and rendered within the animation to illustrate how the techniques presented in this paper can be utilised. As the movie is playing at key points contextual information is displayed in a scrollable pane at the bottom of the interface. A sample of this information is displayed within the figure. The haptic enabled interface permits the software to be utilised using only a visual display and a haptic feedback device.

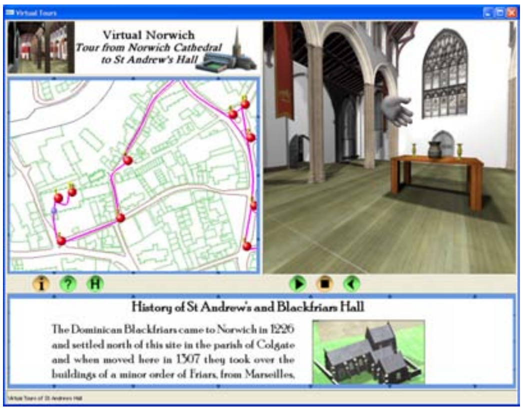


Figure 4: *Image captured during the execution of the program.*

4. Results

To facilitate the construction of test data and to ease the methods applicability, a MaxScript was written to enable all the required information to be extracted from 3ds Max automatically. The user defines any object they wish to be touchable using the "touch" keyword in the object's name and any camera from which to view the scene using the "movie" keyword in the camera's name. By executing the script six cameras for each of the indicated cameras are generated and used to render the six movies for the cube map. All touchable objects are automatically translated to the origin and exported in .3ds format. Furthermore all the touchable object sets and camera and object transformation matrices are exported ready to be used by the C++ implementation.

The fundamental requirement of the proposed system is that it is capable of rendering both the graphics and the haptics at the desired rates of 60Hz and 1000Hz respectively. A high visual quality will be acheived using an offline renderer. To determine if this is possible seven test scenes were constructed using an arrangement of cubes surrounding a camera's path. By executing the script all the necessary models, transformations and movies were automatically obtained and sent to the OpenGL program. The script executed in under one minute for a scene containing four hundred objects. This result excludes the rendering time for the six movies. The simulation was performed and results for the graphics frame rate and haptic feedback rate were obtained, collecting one result every second. Table 1 presents the mean average results recorded. All the tests were conducted on a Dual 3.0GHz Xeon processor computer with 1GB of RAM and a Quadro FX 1440 graphics accelerator card. The haptic feedback is delivered via a Phantom Omni haptic feedback device produced by SensAble Technologies.

The first column of the table states the number of objects included in each test scene. The scenes were constructed by incorporating further objects radially from the camera path's location. Therefore once the scene has increased in size sufficiently the number of touchable objects becomes constant and this number is given in the second column of the table. It can be seen that as the number of touchable objects increases both the haptics and graphics rates remain within an acceptable threshold of the desired goal. Further testing was conducted experimenting with the memory consumed by the movies. It is feasible to consider playing six 50mb movies on each of the faces of the cube at interactive frame rates. However, this is not a crucial factor since one of the systems benefits lies in the ability to construct smaller sections of movie files. This enables a user to change direction more frequently and give them more freedom to explore their surroundings.

# scene objects	# touchable objects	mean fps	mean haptic rate
20	20	604	997
50	50	502	998
100	100	459	998
200	187	295	998
400	279	257	997
800	393	220	998
1600	393	231	994

Table 1: *Comparison of the graphics frame rate and haptic update rate against the number of objects included in the virtual environment.*

5. Acknowledgements

The authors wish to thank the Urban Modelling Group at UEA (http://www.urbanmodellinggroup.co.uk), in particular David Drinkwater for providing photorealistic three dimensional models of St Andrew's Hall, Norwich UK, and its surrounding area. This work was funded by HEART (http://www.heritagecity.org/), acting as a single, co-ordinating organisation to strategically plan, regenerate, manage and promote all heritage resources in Norwich, and Spatial Metro Project (http://www.spatialmetro.org/), a transnational group of partners co-operating to find innovative ways to improve city centres for pedestrians.

5.1. Conclusions

A technique has been presented to enable users to interactively touch large high quality pre-rendered environments. By incorporating a haptic feedback device for navigation and exploration the user obtains a greater level of immersion despite the nature of the semi-prescribed path. Scenes can be constructed in 3ds Max with both dynamic cameras and scene objects and converted into a format suitable for an image based real-time rendering technique automatically by executing a single MaxScript file. This results in all the required information being easily produced for the virtual tour program. The techniques have been illustrated with a virtual tour of St. Andrew's Hall where the user is permitted to touch several objects located inside. Context information is provided to the user both as the movie proceeds and as they interact with objects in the scene.

The current method allows both the camera and the objects to undergo rigid body transformations. For future work techniques enabling the haptic tool to influence the movements of the virtual objects will be investigated.

References

[AKB01] ADAMS R., KLOWDEN D., BLAKE H.: Virtual training for a manual assembly task. *Haptics-e 2(2)* (2001).

[BBF*02] BERGAMASCO M., BROGNI A., FRISOLI A., SALVINI F., VIGNONI M.: Tactual exploration in cultural heritage. In *Proc. of the XIV Computer-aided Egyptology Round Table (IE2002)* (Pisa, Italy, July 2002).

[Bre01] BREWSTER S.: The impact of haptic 'touching' technology on cultural applications. In *Proc. of EVA2001* (Glasgow, UK, 2001).

[CBG00] CROSSAN A., BREWSTER S., GLENDYE A.: A horse ovary palpation simulator for veterinary training. In *Proc. of PHANToM Users Research Symposium* (2000), pp. 79–86.

[Che95] CHEN S.: Quicktime vr: an image-based approach to virtual environment navigation. In *Proc. of ACM Siggraph* (1995), pp. 29–38.

[Che99] CHEN E.: A six degree of freedom haptic system for desktop virtual prototyping applications. In *Proc. of the First International Workshop on Virtual Reality and Prototyping* (1999), pp. 97–106.

[GMS06] GAW D., MORRIS D., SALISBURY K.: Haptically annotated movies: Reaching out and touching the silver screen. In *Demo paper, 14th IEEE Haptics Symposium* (March 2006).

[HBS97] HO C., BASDOGAN C., SRINIVASAN M.: Haptic rendering: Point - and ray - based interactions. In *Proc. of the Second PHANToM Users Group Workshop* (October 1997).

[LD03] LAYCOCK R., DAY A.: Automatically generating large urban environments based on the footprint data of buildings. In *Proc. of ACM Symposium on Solid and Physical Modeling* (2003), pp. 346–351.

[LWR*03] LOSCOS C., WIDENFELD H. R., ROUSSOU M., MEYER A., TECCHIA F., GALLO E., DRETTAKIS G., RECHE A., TSINGOS N., CHRYSANTHOU Y., ROBERT L., BERGAMASCO M., DETTORI A., SOUBRA S.: The create project: Mixed reality for design, education, and cultural heritage with a constructivist approach. In *ISMAR 03* (October 2003).

[MJ06] MORRIS D., JOSHI N.: Hybrid rendering for interactive virtual scenes. In *Stanford University Technical Report CSTR 2006-06* (2006).

[MPT99] MCNEELY W., PUTERBAUGH K., TROY J.: Six degree-of-freedom haptic rendering using voxel sampling. In *Proc. of ACM Siggraph 99* (1999), pp. 401–408.

[OBP05] OWEN R., BUHALIS D., PLETINCKX D.: Visitors' evaluations of icts used in cultural heritage. In *Proc. VAST* (2005), pp. 129 – 136.

[RKK97] RUSPINI D., KOLAROV K., KHATIB O.: The haptic display of complex graphical environments. In *Proc. of ACM Siggraph 97* (Los Angeles, CA, 1997), pp. 345–352.

[RMJ*05] RIZZO A., MCLAUGHLIN M., JUNG Y., PENG W., YEH S., ZHU W.: Virtual therapeutic environments with haptics: An interdisciplinary approach for developing post-stroke rehabilitation. In *Proc. of CPSN05* (2005).

[SMS*01] SNIBBE S., MACLEAN K., SHAW R., RODERICK J., VERPLANK W., SCHEEFF M.: Haptic techniques for media control. In *Proc. of the 14th Annual UIST 2001* (November 2001).

[WSJH04] WILLIAMS R., SRIVASTAVA M., JR. R. C., HOWELL J.: Implementation and evaluation of a haptic playback system. *Haptics-e 3*, 3 (2004).

[ZS95] ZILLES C., SALISBURY J.: A constraint based god-object method for haptic display. In *Proc. of the IEEE Conference on Intelligent Robots and Systems* (1995), pp. 146–151.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Mosaic Rendering using Colored Paper

Yong Jea Gi[1], Young Sup Park[1], Sang Hyun Seo[1] and Kyoung Hyun Yoon[1]

[1]Computer Graphics Laboratory, ChungAng University, Seoul, Korea

Abstract

This paper proposes a new method of simulating colored paper mosaic using computer graphics technologies. This new method focuses on two problems that need to be taken care of in order to simulate colored paper mosaic. The first one is tile generation and the other one is tile arrangement. To get similar result in real art work, we create colored paper object with simple structure. Then, we generate torn colored paper tile by applying voronoi diagram and others to colored paper object. At last, we come up with result images by arranging torn colored paper tile appropriately according to energy function.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: colored paper mosaic, colored paper tile generation, and arrangement

1. Introduction

Mosaic is a technique that uses a pattern or a picture by attaching various tiles such as stone, pieces of broken glass, ceramic pieces, and paper with mortar, lime, cement and glue on a plane. The mosaic technique is much used for painterly decorations in architecture and art works. Furthermore, beyond the application in real life, in computer graphics, there are efforts made to create mosaic images.

This paper proposes a new method of generating colored paper mosaic using computer graphics techniques. To express colored paper mosaic, following two problems should be taken into consideration. First is an algorithm which generates a paper tile that has a torn shape and the second is an algorithm which arranges the paper tile on the appropriate location. In the existing studies on tile mosaic, various tile generation algorithms are proposed. For instance, centroidal voronoi diagram (CVD) [Hau01] and energy-based framework [KF02] are the algorithms that are suggested in order to locate the tile densely in order to arrange the tile appropriately. In contrast, the existing studies on colored paper mosaic ignored the matter of arrangement the paper tile and focused only on the technique of generating paper tiles. For instance, Random method [PKJY00] and random midpoint distribution [SPKY01] are algorithms both proposed for creating torn shaped colored paper tiles. In this paper, we consider the matter of both of tile generation and arrangement for colored paper mosaic.

To generate colored paper tile, we assume that colored paper consists of simple structure (polygon) and material data (texture, color). Then, voronoi diagram [BKOS97] is applied on the colored paper to determine the outline of colored paper tile. And add torn effect by applying random point displacement algorithm onto the determined outline, it finally generates colored paper tiles. Lastly, these generated colored paper tiles are arranged on the optimum location where the calculated energy value according to their location and direction is the largest. This task is repeated until there is no enough space to attach the paper tiles and it finally generates colored paper mosaic.

The tile generation algorithm in this paper has few advantages. First, since the process of generating colored paper tiles is similar to that of tearing the colored paper with hands, it could achieve natural shaped tiles. Second, instead of colored paper, when other types of papers that have other structure or material data such as newspaper and magazine are used, it is easy to get other types of result images.

2. Related work

In the NPR field, there were lots of studies done about mosaic so far. The existing studies on mosaic show impressive results. However, the existing studies only suggest the solution for either the generation or arrangement of the tile.

Hausner [Hau01] proposed a method of arranging tiles similar to the mosaic of ancient Byzantine era using CVD. However, the shape of tiles was limited to squares. To use the mosaic tile that has arbitrary shapes, Kim[Kim'02] suggested energy-based framework. But JIM requires a lot of time in the process of image searching and arrangement in Jigsaw Image Mosaic (JIM). To improve this, Smith [SLK05] applied centroidal area voronoi diagram (CAVD), and Dalal [DKLS06] applied the sum of squared distances (SSD) and fast Fourier transform (FFT). And they also suggested a technique that could create mosaic animation. However, since the existing studies use fixed or pre-determined tiles like square or image, they did not take tile generation into consideration. So, it is hard to directly apply to colored paper mosaic because colored paper tiles are generated dynamically.

The existing studies about colored paper mosaic suggest basic techniques that could generate torn shaped colored paper tiles. Park [PKJY00] applied random method to depict the torn section of colored paper. However, since paper tile is generated too randomly, it resulted in unnatural. To solve this, Seo [SPKY01] suggested random midpoint displacement (RMD) algorithm. RMD is an algorithm generating irregular sections by repeating the displacement of center of the edge perpendicularly. RMD method could enable the generation of natural torn shaped colored paper tile. However, since the existing colored paper mosaic studies did not take the method of arranging colored paper tiles into consideration, the feature lines (edges) of the source image could not be represented accurately. So, Seo [SKPY01] used an image segmentation and quad-tree to maintain the feature lines of a source image. Image segmentation and quad-tree can maintain the feature lines of the source image. but the various shapes of tiles like concave or overlapping of the tiles which is expressed in colored paper mosaic are not represented.

In this paper, we suggest a technique which takes in consideration of the method of generating and arranging colored paper tiles. At first, generate colored paper consists of polygon and material data. Then, tear colored paper. And arrange the generated colored paper tile according to the energy function. In fact, because our algorithm simulate the process of tearing colored paper similar to real, various shapes of paper tile is generated. Also, by adjusting the parameters in the energy functions, we can adjust the degree of overlap and density between colored paper tiles.

3. Preparation

3.1. The structure of colored paper

To simulate the process of real colored paper mosaic, we define the colored paper which consists of two polygon layers like Figure 1. The upper layer represents the color of the colored paper and the lower layer represents the white paper that appears when the colored paper is torn. Also, each layer includes the material data like color and texture image to express the feel of material of paper. We manage these colored papers by establishing color list (L_c) from predefined dataset (in file) of colored paper. The dataset includes the size, texture image, and color of colored paper.

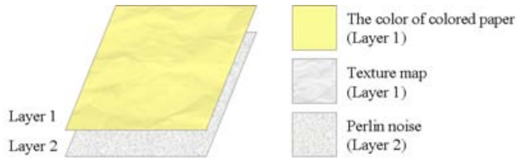


Figure 1: *The structure of colored paper object*

3.2. Source image handling

The inputted source image is segmented into several regions and transformed into the container image. To generate the container image, the image is segmented using mean-shift segmentation [CM02]. However, since it is impossible to extract the feature lines that user expects automatically with mean-shift segmentation. So, The interface is provided the users to merge the segmented regions. Figure 2(b) is an image that the region is merged by the user. Then, extract the feature lines like Figure 2(c), and generate the distance map(D) like Figure 2(d) by caculating the shortest distance between feature lines and each pixel of image. The distance value will be used for calculating energy function to arrange colored paper tiles densely.

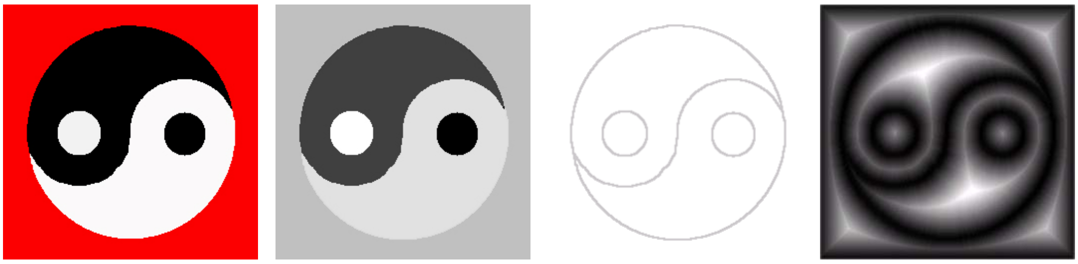
4. Preprocessing

4.1. Initial location of tile

The most basic step in generating mosaic image is to select the initial location which the colored paper tile will be attached. The color and size of tearing colored paper is determined from the initial location. Generally, people attach colored paper tiles based on the bigger pieces from the feature lines of the image. By taking this into consideration, the initial location of the paper tile is defined as location (x,y) in which the $P(x,y)$ is at its maximum. ($D(x,y)$ means distance value of pixel (x,y).)

$$P(x,y) = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} D(x,y) - D(i,j) \cdot [x \neq i \& y \neq j] \quad (1)$$

If $[x \neq i \& y \neq j]$ is equals to $x \neq i$ and $y \neq j$, it is 1, it is not, 0. The term of $D(x,y)$ determines the size of colored paper tile and the term of $D(x,y) - D(i,j)$ determines how to close to the feature lines of image. That is, as each term gets larger, it means that the pixel (x,y) is close to feature lines and can attach bigger tiles.



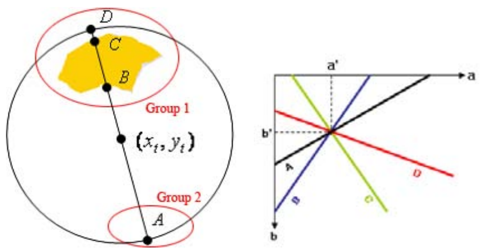
(a) Source image (b) Container image (c) Feature lines (d) Distance map

Figure 2: *Source image handling*

4.2. The color and size of tile

The color of colored paper to each container is determined by average color of the container. First, the threshold value of T_c is determined and select the colors as a candidate which the difference with the container average color is less than T_c from the color list(L_c). Then, one is selected randomly from the candidate on demand.

If the colored paper is selected, determine approximate size of colored paper to be torn. The distance value, $D(x_i, y_i)$, of the initial location (x_i, y_i) refers to the minimum size of colored paper tile that could be attached. So, to attach the tile that is as big as possible, the furthest value $D'(x_i, y_i)$ from (x_i, y_i) is measured by applying Hough transform. First, as in Figure 3(a), the points that have similar gradient from the (x_i, y_i) are grouped into two; one has greater value than y_i and the other group has less value than y_i. Then, in the two groups, select two points that each has the shortest distance from (x_i, y_i) and determine the bigger value between these two points as $D'(x_i, y_i)$. Finally, the size(WXH) of the tile that would be torn is determined as $D(x_i, y_i)XD'(x_i, y_i)$. However, the minimum size(T_{min}) and maximum size(T_{max}) of the colored paper tile is limited by the user.



(a) (b)

Figure 3: *Hough transform*

5. Tile generation

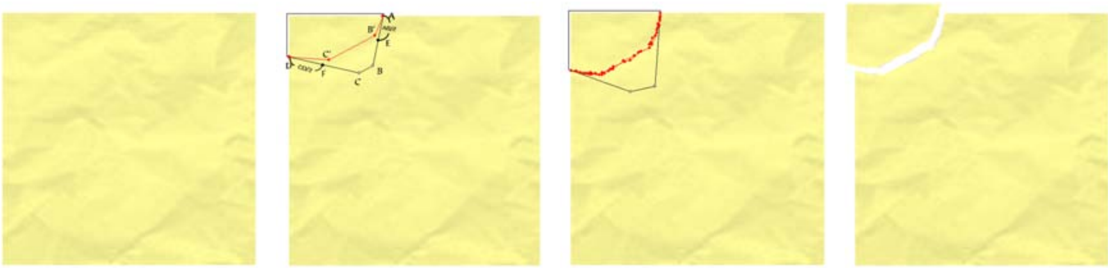
This paper applies voronoi diagram and random point displacement (RPD) algorithm to generate a torn shaped colored paper tile. First, apply voronoi diagram to the colored paper and generate voronoi polygons. Then, randomly select one from the generated polygon placed on the corner and determine as the perimeter of colored paper tile. Next, to achieve more natural colored paper result, smooth the perimeter of the selected polygon and apply torn effect to depict as if the colored paper is actually torn. Finally, tear out the generate polygon from the colored paper to generate colored paper tile. Figure 4 shows the process of tile generation.

5.1. Outline of tile determination

To generate torn colored paper tile, determine the outer perimeter of tearing colored paper tile. We determine the perimeter of the colored paper tile by applying voronoi diagram. First, divide the colored paper tile into the size determined in section 4.2 according to $D(x_i, y_i)XD'(x_i, y_i)$ into many grids. And generate voronoi diagram by placing voronoi sites into each grid. Next, we select which one will be determined for the perimeters of the colored paper tile among the generated voronoi polygons. Generally, people tear colored paper from the perimeter. Therefore, to simulate this process in this paper, determine the polygon including the vertex of the colored paper among the polygons as the outer perimeter of the colored paper tile like black line in Figure 4(b).

5.2. Torn effect

Transform the perimeter to represent as if it is torn colored paper. The polygon generated by voronoi diagram could have rough edges as the polygon depicted in black in Figure 4(b) according to the location of the voronoi sites. When the real person is tearing the colored paper, the rough edges are unnatural since he or she doesn't change the direction of power suddenly when tears a paper. Therefore, in this paper, to create more natural colored paper, apply Bezier curve [Mor99] to smooth out the edges of the polygon. As in Figure 4(b), let the points, A and D as knot vectors and the points E and F that bisect each edge of $\bar{AB}$ and $\bar{DC}$ as control points. However, if the length is shorter than a specific



(a) 2-layerd colored paper object (b) Smooth torn edge (c) Torn effect (d) Tear colored paper tile

Figure 4: *The process of tile generation*

value like $\bar{BC}$, ignore it on the process of generating Bezier curve. This is to create smoother edges represented in red in Figure 4(b). Lastly, transform the smoothed polygon as if the actual colored paper is torn using RPD like red line in Figure 4(c). RPD algorithm represents the irregularity of the torn sections by displacing a random point on the perimeter of the polygon in perpendicular direction. Also, by applying RPD algorithm onto the two layers of colored paper independently, the lower layer, that is the white paper which appears when tearing colored paper could be expressed. Finally, by clipping these modified torn shaped polygons from the colored paper, generate colored paper tile like Figure 4(d).

6. Tile arrangement

When the torn shape colored paper tile is generated, perform the task of arranging this tile onto the appropriate location. The process of arranging colored paper tile is similar to the energy based framework of JIM. However, in contrast, JIM is an algorithm that selects appropriate tiles on a specific location, this paper moves and rotates a specific colored paper tile onto the appropriate location. For each pixel on the perimeter(B) of the tile, let C as container, O as overlapped pixel by another tile, W_e and W_o as weight value, energy function is defined as in equation 2.

$$E = \sum_{i,j \in B} E(i,j) \tag{2}$$

$$E(i,j) = \begin{cases} -D(i,j) \cdot W_e & (i,j) \notin C \\ -D(i,j) \cdot W_o & (i,h) \in O \\ T_{max}/2 - D(i,j) & \text{else} \end{cases}$$

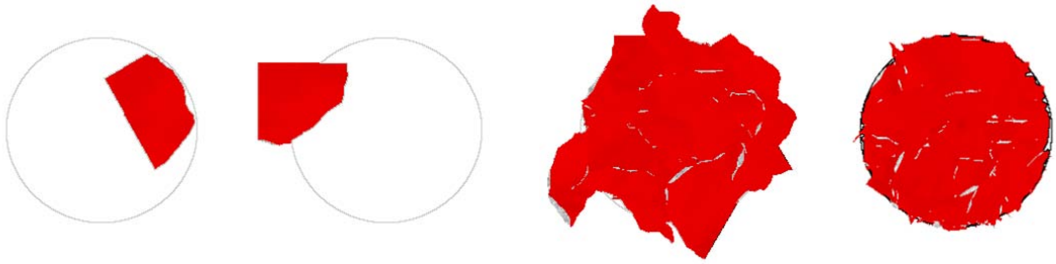
The most appropriate location for attaching colored paper tile is the location which the total of energy function E is at the biggest. The optimum location will be searched until there is no bigger E value in the neighboring pixel starting from the initial location(x_i, y_i). The energy value E could generate various results according to weight value. When W_o is larger than W_e, the tiles are less overlapped like Figure 5(c), and when W_e is larger than W_o, the feature lines of the source image could be maintained like Figure 5(d). Figure 5 shows the result of arrangement according to the weight value. Mostly, it finds the location and direction appropriately as in Figure 5(a), but some times, there are instances where it is inappropriately arranged as in Figure 5(b). In these instances, move the tile inside the container region and apply energy function. Since this algorithm is the brute-force method, it takes a long time in searching for the optimum location. This is a matter to be taken into consideration in the future.

7. Results

All of the final images in this paper are created by setting the value of W_e as 5 and W_o as 2. Figure 6 shows various results depending on the smallest value(T_{min}) of colored paper tile. If the smallest value of tile gets bigger, more grout space will be created so the result image will come out incorrect. This difference is well shown in Figure 6(a) and Figure 6(b). However, if the smallest value of tile is set to be too small to avoid the problems mentioned above, then too many small tiles will be arranged in places so the texture of tile will not be expressed well. Therefore, it is necessary to set appropriate T_{min} value(5 or 6 pixels) of tile in order to get desired result image. And there is unexpected problem occurred in Figure 6(b). It is because colored paper tiles are created in arbitrary shape so tile bigger than container can be generated sometimes. This problem can be solved by limiting the maximum value(T_{max}) of tile in each container region independently. However, the best solution is generating the shape of tile user expected by adjusting the location of voronoi sites. Figure 7, Figure 8 and Figure 9 show another rendering results. Top left portion of result image is source and container image. It took about 40-60 minutes to create result image.

8. Discussion and future work

This paper suggests the method of considering the generation and arrangement of the tile to simulate colored paper mosaic. Through colored paper generation algorithm, more



(a) The best case (b) The worst case (c) $W_e = 1, W_o = 5$ (d) $W_e = 5, W_o = 1$

Figure 5: *Comparison arrangement algorithm according to weight value*

natural and various shaped tiles are generated than the existing studies. And by adjusting the weight value and other parameters of energy function, we can arrange the colored paper tiles to be overlapped or not. The method of this paper is to simulate as the actual colored paper mosaic to achieve more natural colored paper mosaic effect.
However, this paper has some points to solve. First, since the shape of paper tiles are randomly generated, there is a possibility of tile generation that would be larger than the container. To solve this, the method of arranging voronoi sites to generate the tile in the shape that the user intended to is necessary. Second, since our algorithm use only texture image for feel of material, the result is unnatural. So, some methods of representing feel of material to simulate actual colored paper mosaic are necessary. Third, since the algorithm of arranging the tile in this paper is brute-force method, it takes a long time. The algorithm such as FFT or hardware supporting such as GPU should be considered for performance. Finally, if the mosaic image is generated by applying objects that have other characteristics instead of colored paper object, other impressive results could be achieved.

(a) T_{min}=10, T_{max}=80 (b) T_{min}=5, T_{max}=80

Figure 6: *Comparison results according to T_{min} value*

9. Acknowledgement

This work was supported by grant No. (R01-2005-000-10940-0) from the Basic Research Program of the Korea Science & Engineering Foundation.

References

[BKOS97] BERG M., KERVELD M., OVERMARS M., SCHWARZKOPF O.: *Computational Geometry Algorithms and Applications*. Springer, 1997.

[CM02] COMANICU D., MEER P.: Mean shift: A robust approach toward feature space analysis. In *IEEE Trans. Pattern Anal. Machine Intell.* (2002), pp. 603–619.

[DKLS06] DALAL K., KLEIN A., LIU Y., SMITH K.: A spectral approach to npr packing. In *NPAR'06: Proceedings of the 4th International Symposium on Non-Photorealistic Animation and Rendering* (2006), ACM, pp. –.

[Hau01] HAUSNER A.: Simulating decorative mosaics. In *Proc. SIGGRAPH '01* (2001), vol. 20, pp. 573–580.

[KF02] KIM J., F.PELLACINI: Jigsaw image mosaics. In *Proc. SIGGRAPH '02* (2002), vol. 21, pp. 657–664.

[Mor99] MORTENSON M.: *Mathematics for Computer Graphics Applications 2nd Edition*. Industrial Press, 1999.

[PKJY00] PARK Y., KIM Y., JHO C., YOON K.: Mosaic techniques using color paper. In *Proc. KCGS '00* (2000), pp. 42–47.

[SKPY01] SEO S., KANG D., PARK Y., YOON K.: Colored paper mosaic rendering based on image segmentation. In *Proc. KCGS '2001 Conference* (2001), pp. 27–34.

[SLK05] SMITH K., LIU Y., KLEIN A.: Animosaics. In *Computer Animation 2005 : ACM SIGGRAPH / Eurographics Symposium on Computer Animation* (2005), ACM, ACM, pp. 201–208.

[SPKY01] SEO S., PARK Y., KIM S., YOON K.: Colored paper mosaic rendering. In *Proc. SIGGRAPH '2001 Conference Sketches and Abstracts and Applications* (2001), p. 157.

Figure 7: *Colored paper mosaic rendering (1024X768, $T_{min} = 6$, $T_{max} = 40$)*

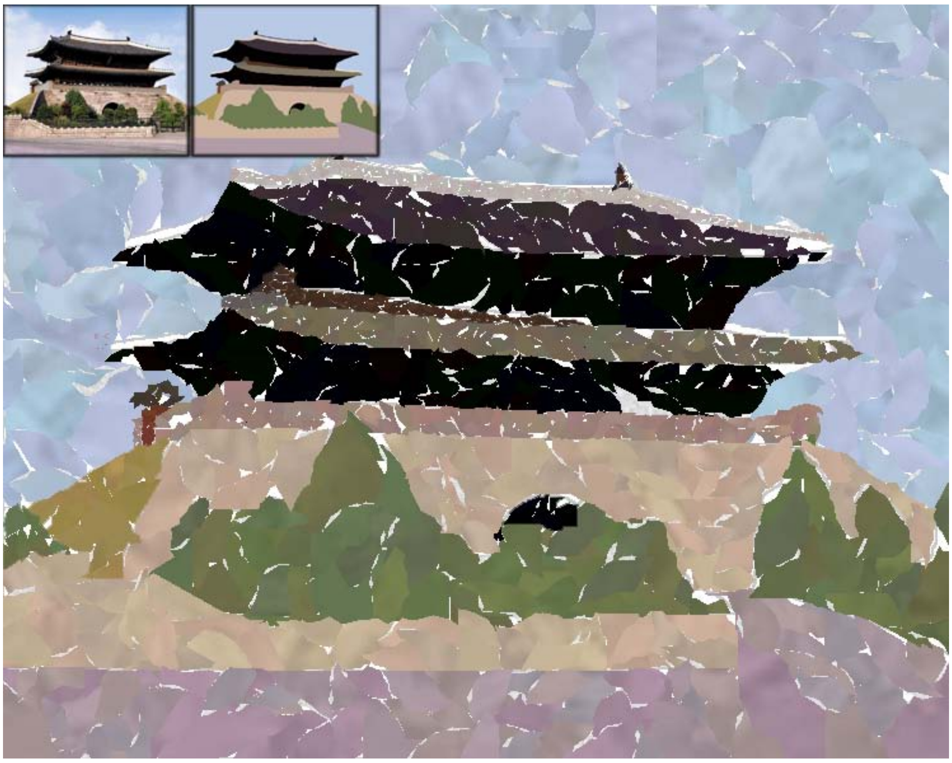


Figure 8: *Colored paper mosaic rendering (750X595, $T_{min} = 5$, $T_{max} = 100$)*

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Restoration of Color in Noh Masks Based on Retinex Algorithm

Xin Yin[1], Yasuo Miichi[2], and Hiromi T.Tanaka[1]

[1]Ritsumeikan University, Japan
[2]Noh Mask Maker, Japan

Abstract
Noh is a famous traditional Japanese dramatic art. The restoration of ancient Noh masks is an important aspect of the preservation of cultural heritage. In this paper, we developed a technique based on Retinex algorithm and it can be applied to restore the color of 3D objects such as the Noh mask. The Retinex algorithm simulates the ability of human eye and can adjust the color of the 2D image. To apply the Retinex algorithm on the 3D surface, the shape of the object is taken into account. The original color of the target points can be got by reference to the color of the surrounding points sampled according to a uniform route. The effectiveness of this technique is demonstrated by the implementation results. This technique can be used to infer the color of other 3D objects also.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Color, shading, shadowing, and texture

1. Introduction

Objects of cultural assets which are exposed to the natural environment can undergo changes in appearance. The original colors may fade and some parts of the object may be lost. For this reason, much research has been undertaken into the preservation and restoration of cultural assets. Various digital restoration techniques have been developed for preservation and exhibition of cultural assets. In this paper, we introduce a technique for color restoration in Noh masks.

Noh is a famous Japanese traditional drama in which the faces of the players are covered with masks. One Noh mask is shown in the Figure 1. Noh masks are vulnerable to environmental weathering which can result in fading and loss of the surface pigments. The restoration of color in Noh masks is a challenging area of research.

Figure 1: *The photo of Noh mask.*

2. Previous Work

Long ago, the restoration of cultural objects was based on the rich knowledge and experience of craftsmen. More recently, digital techniques using image processing software have been developed, but like the craftsman of former times, the operator applies his or her knowledge and experience in restoring the color of the object. One example of this type of technique is [CN02], by which the color of an ancient drawing was restored perfectly. In addition, some color restoration techniques based on computer graphics (CG), such as [YT04], have been developed. In this technique, the three-dimensional (3D) shape of the object is measured by a machine. Next, the artist draws pictures based on the historical record, which are mapped onto the surface of the 3D model by computer. The color of an ancient statue was reproduced using this technique. This technique allows manual adjustment of detail based on the experience of the artist. But these techniques pain the color on the surface of object and lost some real information of color on the surface.

Automatic or semi-automatic color restoration techniques have also been developed. [SBJ05] used a technique which makes use of color information remaining on the object. A color that is lost may be reproduced by reference to the color information surrounding the missing section. [LLP00] developed a system for restoring the color base on the experience of the exporters. [BBC00] introduce some digital techniques which can clean the surface and repair the crack pattern on the objects. And [DC05] developed an algorithm, based on the ability of human eyes to adjust to environmental lighting conditions, for removing the influence of dirt on computer simulations of ancient drawings. This technique has mainly been used in the color restoration of oil paintings. These techniques mainly restore the color on 2D images. We propose a technique for restoring the color on 3D objects in this paper.

Usually, for restoring the color in the culture assets, the investigation of the pigment on the surface and segmenting the pigment field is needed. But there are some problems are not solved well. The same pigment can show different color if the pain method is different. So it is very difficult to understand original color in the culture assets even we know what pigment is used. By other hand, segmenting the pigment field well is not easy. The segment technique is developing now. To avoid these problems, we develop a technique to restore the color based on the information of the pigment on the surface directly. The color variation process is different. The craftsmen always restore the color by reference the color where is look new on the surface. This means that the different color on the surface implies the some rules of color variation process. If using the different color information on the surface well, the original color can be restored. The technique represented in this paper needs not segmenting the color field. The color of the target points is updated by the color of the surrounding sample points. Shown as Figur 2, The first step was measurement of the mesh and the color of the Noh mask. Next, we sample the variation in color over the 3D surface based on the surface normal variation and the color of the Noh mask was restored based on the Retinex (come for the retina and cortex) algorithm. Finally we show some implementation results and discuss this technique.

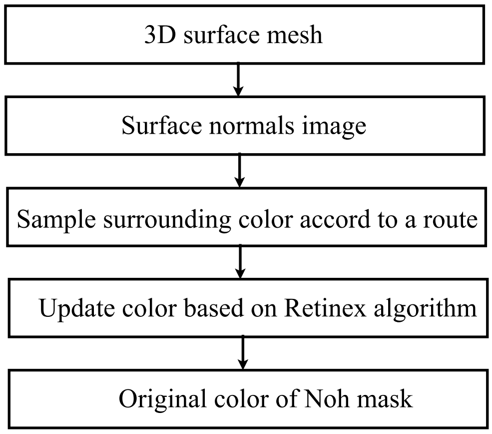
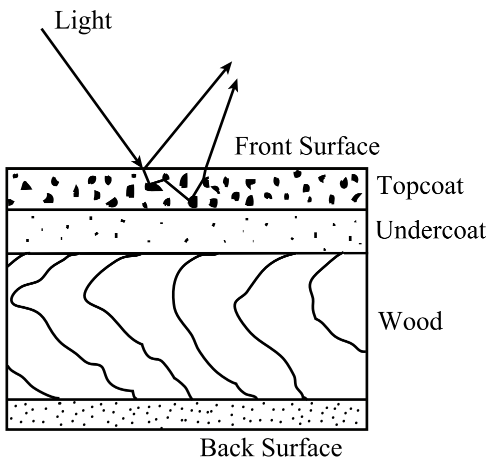


Figure 2: *Compute processing.*

3. Background Knowledge on Noh masks

Noh masks are made from wood on which a design is pained([Iha02]). In creating a Noh mask, the first step is to carve and polish the wood to provide detail, so that a shape similar to that of the human face is obtained. Next, pigments are painted onto the wooden mask. As shown in Figure 3, the Noh mask has two surfaces, the back and the front. The back surface is nearest to the face of the player, and is painted uniformly because it cannot be seen. The front surface is painted with various types of pigments. There are two types of coating: the undercoat and the topcoat. The undercoat is painted in a uniform color and represents the color of the skin. The topcoat is painted in different colors to show the hair, eyes, mouth and so on. Finally, certain other pigments are applied to the surface so that the Noh mask can be seen well on the stage.



There are two pigment coats on the front surface and one pigment coat on the back surface.

Figure 3: *Pigment of Noh mask.*

A glue is used to cause the pigments to adhere to the wood. Light passes through the glue and is reflected outward by pigment particles. When the glue deteriorates, some of the pigment particles are lost, with the result that the formerly smooth surface of the mask becomes rough and decreases in brightness. If the deterioration extends deep into the mask, some areas of pigment are lost completely, and the color of the mask cannot be seen. Since the pigment particles are mixed slowly, the color on surface look darker. One subject of restoring the color is to take off this dark efficiency.

4. Restoration of Color

The techniques used in ancient times for fashioning and painting Noh masks are not well understood. In addition, the variation of pigments is a complex process. For these reasons, it is very difficult to restore color completely. The key idea for restoring the color in this paper come from the [DC05] which use Retinex algorithm ([LM71]) to restore the color on 2D image. To update the color of a target point, the Retinex algorithm uses the color information found in the vicinity of the point. The influence of one pixel on another varies according to the distance between them. Similar to the Retinex algorithm, we also use the color information found in the vicinity. Then compute the color based on this information. The detail will be introduced in this section.

4.1. Measure the the Color and 3D Mesh

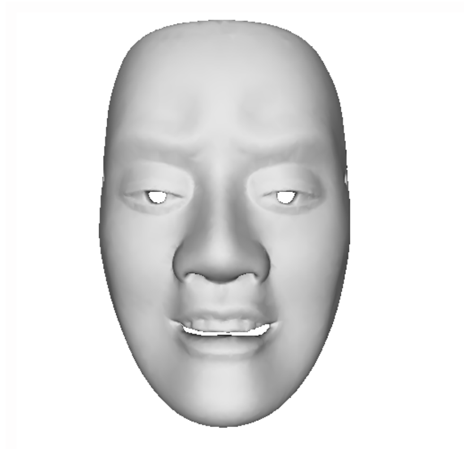
Because the Noh mask is a 3D object, the normal of the 3D object can cause variations when the color of the mask is measured. The same color may show a different value if the view point is different. For example, there are some highlight and shadow when take photo of a mask shown as Figure 1. To avoid this highlight effect, the mask is put into a white dome. In this white dome, the light is diffused and there is not highlight on the mask when take the mask photo. Use this method, we can get the initial color image of Noh mask without highlight.

The 3D surface data were obtained using a 3D object measurement machine (Vivid 910). Using exist software to compose these data together, the 3D mesh surface shown as Figure 4 is got. The 3D mesh will be used to get the image which show the normal of the surface and the image which show the vector from surface to camera. These image is useful to decide the sample position when renew the color on the surface late.

Since the mesh is 3D and the color image is 2D, the registration process between the 3D mesh and the 2D image is necessary. Three reference points are used. Two points are the center of the eyes. One point is the center of the mouth. Then the 3D mesh is rotated and scaled until these three points are match each other. After this process, the mesh is look from same viewpoint with the 2D photo of the Noh mask.

4.2. Sample on 3D Surface

As mentioned before, when update a point (called target point) color, the surrounding point (called sample point) color information is reference. To get uniform samples, a sample route shown in Figure 5 is used. The color of point A is updated based on the colors of the surrounding points (circle points). [YS05] has introduced the concept of aesthetic curves. We assume that these curves are beautiful because color information obtained as along with these curves are uniform. To decrease the compute cost, the expression of the curves is simplify and is represented in polar coordinates as follow:



This surface model is constructed from about 20 different patches and compose with 20,000 triangles.

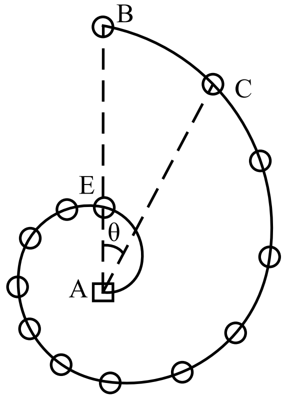
Figure 4: *3D mesh of Noh mask.*

$$\begin{cases} R_2 = a \cdot R_1 \\ \theta = b \end{cases} . \qquad (1)$$

Where R is the length between the point A and sample points. For example, R_1 is the length of AB and the R_2 is the length of AC. Shown as Figure 5, θ is the variation in angle. a and b are constants. In this paper, a is 0.9 and b is $\pi/6$. The surrounding points (circle points) are distributed according to this curve, and the color of point A is calculated from te color of the surrounding points (B, C etc.).

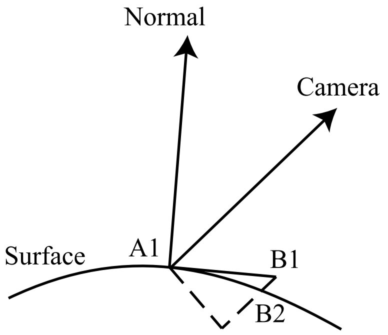
But the surface of the object is 3D. It is necessary to adjust the sample method on the surface according to the normal of the surface. Shown as in Figure 6, A1 is the point which needs updating the color. B1 is the sample point which is compute by the exoression (1). B2 is the projection of B1 on the surface according to the vector to the camera. As the result, the sample point on the 3D surface is the B2. For decide this new position, there are parameters connect to the shape of object are used. One is the normal of surface which can be used to compute position of B1. Another is the vector from surface to the camera and is used to decide the position of B2.

For generating the route, the normal information and the vector to the camera information is draw into the normals



The color of target point (point A) is updated by the color of surrounding sample points (point B etc.) according to this uniform route.

Figure 5: *The sample route.*



The sample position is adjusted from B1 to B2 parallel with the vector of camera.

Figure 6: *Adjustment of the sample route.*

image and the camera-vector image shown as Figure 7. The values of the RGB color show the xyz values of the vectors. By this translation, the 3D problem is translated to the 2D problem. As the result, the process of restoring the color becomes simple. The sample route can be adjusted from these vector images.In the normals image, three states of the route on the different position of the surface are shown. Then according to these routes, the color of the target points is updated.

4.3. Update the Color

This color updating is calculated in the logarithmic field using the LMS color system. RGB color values are converted into LMS values in the logarithmic field, and the color of each point is then updated along the new route. Finally, the LMS color values in the logarithmic are converted back to the RGB system. The discussion at next is in the LMS color system.

The speed of the pigment variation on surface is different,so these different information can be used to restore the original color of the mask. Usually, the process of the pigment variation is similar on the different superficial position. If the position is near, the pigment variation is more similar. As this reason, when update the color of a target point. The color of the sample points along with the route will give some contribution to the color of the target point. The near points give more contribution. Based on this idea, the next expression is constructed:

$$NP = OP + \sum_{i=1}^{N}((OP_i' - OP)/(k^{(N-i+1)})) \ . \qquad (2)$$

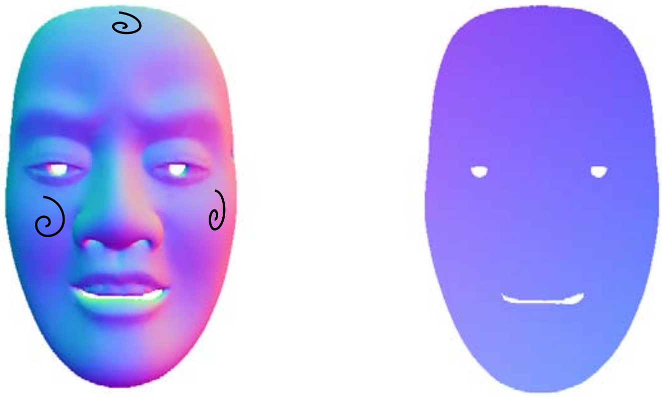
Here, NP is the new color of the target point; OP is the old color of the target point; OP_i' is the old color of the sample points (points B and C in Figure 5). i is the index of the sample points along with the sample route. N is the total number of the sample points and is 12 in this paper. If the index i is small, the distance between the sample point and the target point is large. k is a constant and is the parameter that control the amount of influence one point has on another. In this paper, k is 2.0. This update process is done by some times. If the same pigment shows the same color, it is mean that the original color is got and the compute process is finished.

Despite the pigment variation process is not understood well, the difference of the color on surface can help us to know something about the pigment variation. The compute process that mentioned above can infer the original color of the pigment automatically.

5. Implementation and Discussion

Implementation is carried out using two Noh masks. One was made in about thirty years ago. The other mask is famous one that was made in about three hundreds years ago. Shown as in Figure 8, the current mask (top left image) color is a little dark. After the cleaning process, the color becomes new and is bright (top right image). This mask is made in thirty years ago. The maker of the mask says that the color is similar to the original one in his memory. Another mask was made about three hundreds years ago and nobody know the original color. At this case, the photo is projected to a column surface and a 2D texture is got. The sample algorithm and cleaning process is same as above. And then the result texture is mapped back on the 3D mesh surface. The color of the accient Noh mask before restoration (bottom left image) is dark. After the restoration process (bottom right image), all of the color is bright and look like a new one. Especially, the color of the skin and the red color of the lip is look well.

The Retinex algoritm can decrease the same color component such as light effect.The dirty on surface is a same color componet also, so the Retinex algorithm is successful to remove the uniform dirty effect on the surface. In future, it is



The value of RGB color represent the value of xyz of the surface normal (left image) and the camera vector (right image).These two image is used to adjust the sample route on the 3D surface.The black curves in normal image show the route adjusted.

Figure 7: *Normal image and camera-vetor image.*

need to evaluate the result by scientific method. The technique in this paper is a digital technique to restore the color of the Noh mask. It is not consider the weathering process of the pigments of the Noh mask. And because the classes of the ancient Noh pigments is different from that used today, It is need to investigate the weathering of these classes pigments in future. In some cases, the pigment of Noh mask is lost. The technique presented in this paper can not restore the original color of Noh mask in this case. So it is need to analysis the original pigment types and restore it by developing a new technique. This technique need not segment the pigment field and can be applied on the 3D object easily. These are the main advantages of this technique.

6. Summary

In this paper, we describe the color restoration of a Noh mask by use of a sample route which takes into account the 3D shape of the object. The effectiveness of this technique is demonstrated by the results obtained. The technique can also be used to restore color in other 3D objects. To further develop this work, we intend to calculate the influence of weathering on the pigment colors in order to obtain a better result. We will also attempt to develop a new rendering technique which takes into account the pigment microstructure of the Noh mask.

Acknowledgments

This work was supported partly by <Kyoto Art Entertainment Innovation Research> of Centre of Excellence Program for the 21st Century by Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[BBC00] BARNI M., BARTOLINI F., CAPPELLINI V.: Image processing for virtual restoration of artworks. *IEEE MultiMedia 7*, 2 (2000), 34–37.

[CN02] CO.LTD D. D., NHK: Restoration the color of hell scroll. http://www.dnp.co.jp/nenshi/nenrin/lab/syufuku.html (in Japanese).

[DC05] DRAGO F., CHIBA N.: Locally adaptive chromatic restoration of digitally acquired paintings. *The International Journal of Image and Graphics 5*, 3 (2005), 617–637.

[Iha02] IHARA S.: Seigen ihara and noh mask. http://www5.ocn.ne.jp/~seigen/ (in Japanese).

[LLP00] LI X., LU D., PAN Y.: Color restoration and image retrieval for dunhuang fresco preservation. *IEEE MultiMedia 7*, 2 (2000), 38–42.

[LM71] LAND E., MCCANN J.: Lightness and the retinex theory. *Journal of the Optical Society of America 61*, 1 (1971), 1–11.

[SBJ05] SCHAR S., BIERI H., JIANG X.: Digital restoration of medieval tapestries. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST 2005* (2005), pp. 235–240.

[YS05] YOSHIDA N., SAITO T.: Aesthetic curve segment. *IPSJ SIG Technical Report 2005*, 2005-CG-121 (2005), 97–102. (in Japanese).

[YT04] YAMADA O., TAKASE Y.: A case study for the practical use of 3d digital archive of cultural properties. *IPSJ Symposium Series 2004*, 17 (2004), 31–38. (in Japanese).

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Real-time Animation of Various Flame Shapes

F. Bridault[1] F. Rousselle[1] C. Renaud[1] M. Leblond[1]

[1]Laboratoire d'Informatique du Littoral

Abstract

Working on the computer reconstruction of the Gallo-Roman forum of Bavay, we try to improve the feeling of immersion in the virtual environment. One way to achieve this is to provide realistic and dynamic light sources. In this context, we need to model candles, oil lamps, torches or bonfires. We propose in this paper a model that can handle complex flames in real-time and manage interactivity. The fire is considered as a set of linear flames whose shapes are defined by the geometry of the combustible and the fuel distribution. Each individual flame is represented by a textured NURBS surface. Then, combining several real-time effects such as glow and true transparency, we are able to make the NURBS surfaces merge in a convincing way, and to give the impression of a real fire.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computing Methodologies]: Computer Graphics Three-Dimensional Graphics and Realism

1. Introduction

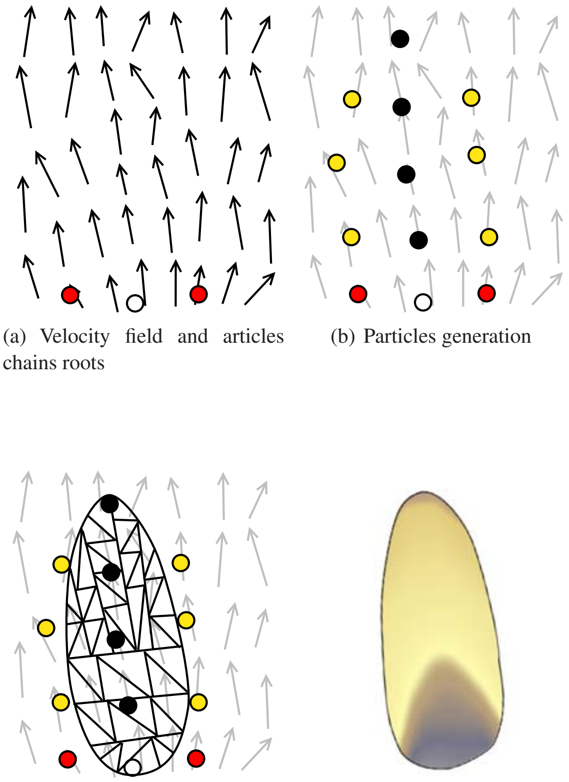
Fire is a challenging problem in computer graphics. Describing flames is of primary interest for movie-making, animated films and video games. This can also be important for virtual environments, especially when these are reconstructions of ancient sites. Indeed, fire was the only source of light until the end of the 19th century. The lack of fire in such virtual environments would decrease realism, and would reduce the possibility of representing scenes in the night or allowing real-time walkthroughs in dark rooms.

An example of such application is the *CyberForum* project which aims to reconstruct and to visualize interactively the Gallo-Roman forum of *Bagacum* (Bavay) in the north of France. This building was built between Flavian (70 A.C.) and Severian (beginning of 3rd century) periods and its impressive dimensions (240 by 110 meters) make this forum the largest known in Gaul to this day. Due to the damage caused over time it is often difficult for any visitor to understand the architectural design of the site. The virtual reconstruction of the different buildings of the forum was thus performed two years ago and is extensively used for visitors through a 3D stereo interactive movie (figure 1). It is also employed by archaeologists as a research tool in order to understand some parts of the site or to compare some of their hypotheses. They are therefore interested in visualizing some parts of the forum lit by flames in order to make better investigations, as it was done in Knossos [IC03]. For the visitors, the addition of real-time flames can improve the feeling of immersion in the virtual site.

Figure 1: *Interactive 3D stereoscopic movie and projection room (courtesy of musée/site archéologique départemental de Bavay; A.Solé 2006)*

In this paper, we introduce an extension of a real-time flame model [BLLRR06] that can handle complex flames thanks to the use of multiple NURBS surfaces. We also present a new intuitive and interactive way to build flames with *virtual wicks* and to model the fuel distribution. Last, we outline the combination of existing multiple real-time rendering techniques to get a realistic appearance.

After words the paper is organized as follows. We first talk about previous work in section 2. Then, section 3 summarizes the mechanism of the flame model we proposed in [BLLRR06], which is the basis of our method. We describe extensively our approach to the modeling of complex flames in section 4, while section 5 presents the techniques we used for the rendering. Section 6 shows our current results and we finally give and discuss some perspectives in section 7.



(a) Velocity field and articles chains roots (b) Particles generation (c) NURBS surface generation (d) Texture mapping

Figure 2: *Mechanism of our previous model for a candle (side view)*

2. Previous Work

The literature contains many techniques to display flames on computers, but to our knowledge, none was really suitable for real-time virtual environments. The first kind of technique uses physics to describe the combustion process. This was done first with simple laminar flames, first static [Ina90], then dynamic [Rac96]. Diffusion equation [SF95] could also be employed, although it implies a high number of parameters, to describe larger turbulent fire. Nguyen *et al.* [NFJ02] use two incompressible Euler equations in addition to a method known as the level-set equation to build an implicit surface representing the blue core. Rendering is done according to the model of the black-body radiation which leads to a very realistic simulation. Lamorlette *et al.* [LF02] also published a realistic model designed for production environments. This model is also interesting since it is able to describe a wide range of flame types, from a simple candle to a dragon's breath. The obvious drawback for all these methods is the computation time, making them unsuitable for our real-time constraint.

Other methods therefore try to avoid physics and make approximations in order to obtain a better computation time. Particle systems [Ree83] are widely used. They make it possible to describe fuzzy objects with a low computing cost. Beaudouin *et al.* [BPP01] use particle chains evolving in a velocity field and describe a single flame using a potential equation. Pszczolkowska [Psz04] also applies a similar method, adding Perlin's noise [Per85] to model turbulence effects, and a Gauss function instead of a potential equation. Nevertheless, neither of them reach real-time frame rate.

Wei *et al.* [WLMK02] introduced a GPU implementation of the Lattice Boltzmann Model to model air flows. They release particles in the velocity field and then render the fire with textured splats. This model is fast and is able to achieve true real-time simulation and animation, however the fire looks more like a burned gas cluster than real flames.

Rather than simulating the flames, some approaches capture real flames data. In the context of photo-realistic illumination by flames for virtual heritage, Chalmers *et al.* [DC01, Cha02] measure a candle flame illumination using a spectroradiometer. A video-captured flame is then incorporated into a virtual scene and its illumination is computed by approximating the flame shape with several emitting spheres. Later, Hasinoff and Kutulakos [HK03] developed an approach to reconstruct a 3D flame from two 2D orthogonal views. Recently Ihrke and Magnor [IM04] presented a tomographic method for reconstructing a volumetric model of flames from multiple camera images. Although these models are of high quality, they are of course restricted to non interactive animations because they rely on a static set of real images.

3. A simple model

Considering no existing model fits the needs for virtual environments, a real-time model for flames was introduced in our previous work [BLLRR06]. Our goals were to simulate accurately the dynamic of simple flames and to render them realistically in real-time. We used a Navier-Stokes equations solver based on the well-known implementation of Stam [Sta99, Sta03] to create a velocity field. External forces were added at the bottom of the grid to simulate the buoyancy. Particle chains were generated in the velocity field from generation points called *roots* (figure 2(a)). Two kinds of chains were created : the *lead skeletons* and the *peripheral skeletons* (figure 2(b)). The former define the vertical boundaries and the latter the horizontal boundaries of the flame. The particles from these chains were used as control points to build a NURBS surface which approximates the shape of the flame (figure 2(c)). Last, a transparent 2D texture was mapped onto the surface to render the colors of the flame (figure 2(d)).

This method was quite fast and yet had several limitations. Above all, it was restricted to small flames and could not produce complex flames such as torches or camp fires. Then it lacked a simple way to handle the buoyancy of the flames. Forces were defined at the bottom of the flame but there was no understandable model to describe them, although it is crucial to have one, because in this way a simple interface can be defined for the potential users of such a technique. Last, rendering of the flames could be improved.

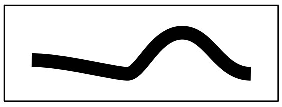
4. Towards complex flame modeling

We introduce several concepts here. First we consider a *virtual wick* for a linear flame. Then we associate a *Fuel Distribution Function* (FDF) with the wick in order to compute the buoyancy of the flame. That being done, we generate a NURBS surface for the wick in the velocity field. Last, to build a more complex fire such as torches or bonfires, we consider a set of flames.

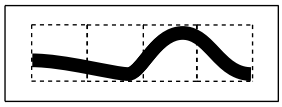
4.1. The virtual wick

As observed during our experiments, the shape of the wick has an impact on the final shape of the flame, that's why it is important to represent it. Actually, the wicks we define should be considered as an interactive tool for the user to place fire fronts and to draw somehow a flame. Indeed, it will be shown later that many wicks are employed to define larger flames although this does not really correspond to reality. Therefore the wicks themselves can be displayed or not according to the type of fire being rendered. The wicks we use are generally long cylindrical shapes as seen in figure 3(a). They can be modeled in any 3D modeler and then imported in our software to be used as a source of fire.

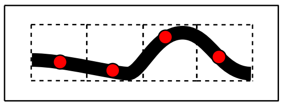
Whereas the roots of the particles chains were placed by hand in our previous work, we here have to set a method to place them automatically on the wicks. The wick in bounding boxes is divided according to the number n_l of lead skeletons we want (figure 3(b)). Generally, two or three lead skeletons are used per flame. Then we compute the barycenter of the vertices in each bounding boxes. This is where the roots of the lead skeletons are placed (figure 3(c)). In addition, two extra lead skeletons roots are added on each extremity of the wick, so that there is a total of $N_l = n_l + 2$ lead skeletons (figures 3(d) and 4(a))). After that, two roots of peripheral skeletons are placed on each side of a lead



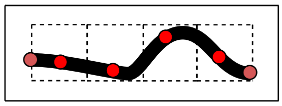
(a) Wick to be processed



(b) Cutting into bounding boxes

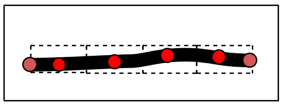


(c) Adding of lead skeletons roots

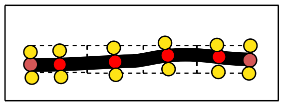


(d) Extra lead skeletons roots

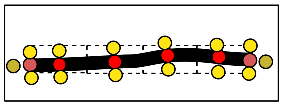
Figure 3: *Placement of the roots of the lead skeletons on the wick (side view)*



(a) Previous wick with roots of lead skeletons



(b) Adding of peripheral skeletons roots



(c) Extra peripheral skeletons roots

Figure 4: *Placement of the roots of the peripheral skeletons on the wick (top view)*

skeleton root, on the wick sides (figure 4(b)). Two extra peripheral skeletons roots are also added at the extremities of the wick (figure 4(c)). This way, there is therefore a total of $N_p = N_l \times 2 + 2$ peripheral skeletons.

4.2. Fuel Distribution Function

The shape of the wick is not sufficient to describe the shape of the flame; fuel is in fact one of the most important factors. In the real world, wood, oil or wax don't produce the same kind of fire. But we can also notice that the fuel is not spread uniformly on the burning surface. In oil lamps, the capillary action in the wick is the main factor. However our real-time constraint does not permit us to perform complex computations. Therefore, we need a simple process to model the fuel distribution.

Buoyancy is simulated by adding external forces in the velocity field of the solver. A simple function $F(u_i)$ is used to describe their distribution. Each root of the lead skeletons is associated with a value u_i in the range $[-1;1]$ or $[0;1]$ according to the type of function. Considering the wick in one dimension, the value of the root *on the extreme left* u_0, is set to -1 or 0 and *on the extreme right* u_{n-1} is set to 1. The other values $u_1 ... u_{n-2}$ are linearly interpolated.

$$A = \begin{pmatrix}
L_{0,0} & P_{0,0} & P_{1,0} & \cdots & P_{M-2,0} & P_{M-1,0} & L_{M-1,0} \\
L_{0,0} & P_{0,1} & P_{1,1} & \cdots & P_{M-2,1} & P_{M-1,1} & L_{M-1,0} \\
L_{0,1} & P_{0,2} & P_{1,2} & \cdots & P_{M-2,2} & P_{M-1,2} & L_{M-1,1} \\
L_{0,2} & P_{0,3} & P_{1,3} & \cdots & P_{M-2,3} & P_{M-1,3} & L_{M-1,2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
L_{0,N_l-2} & P_{0,N_p/2-3} & P_{1,N_p/2-3} & \cdots & P_{M-2,N_p/2-3} & P_{M-1,N_p/2-3} & L_{M-1,N_l-2} \\
L_{0,N_l-1} & P_{0,N_p/2-2} & P_{1,N_p/2-2} & \cdots & P_{M-2,N_p/2-2} & P_{M-1,N_p/2-2} & L_{M-1,N_l-1} \\
L_{0,N_l-1} & P_{0,N_p/2-1} & P_{1,N_p/2-1} & \cdots & P_{M-2,N_p/2-1} & P_{M-1,N_p/2-1} & L_{M-1,N_l-1} \\
L_{0,N_l-1} & P_{0,N_p/2} & P_{1,N_p/2} & \cdots & P_{M-2,N_p/2} & P_{M-1,N_p/2} & L_{M-1,N_l-1} \\
L_{0,N_l-2} & P_{0,N_p/2+1} & P_{1,N_p/2+1} & \cdots & P_{M-2,N_p/2+1} & P_{M-1,N_p/2+1} & L_{M-1,N_l-2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
L_{0,2} & P_{0,N_p-3} & P_{1,N_p-3} & \cdots & P_{M-2,N_p-3} & P_{M-1,N_p-3} & L_{M-1,2} \\
L_{0,1} & P_{0,N_p-2} & P_{1,N_p-2} & \cdots & P_{M-2,N_p-2} & P_{M-1,N_p-2} & L_{M-1,1} \\
L_{0,0} & P_{0,N_p-1} & P_{1,N_p-1} & \cdots & P_{M-2,N_p-1} & P_{M-1,N_p-1} & L_{M-1,0}
\end{pmatrix} \quad (1)$$

The fuel distribution function $F(u_i)$ itself can be any of the following : linear, bilinear, exponential, Gaussian, random. Of course it can also be any user-defined function. When we add the forces corresponding to the buoyancy of the flame, we put a vertical force in the voxel of the solver grid where each root is placed. The force f_i is finally computed as :

$$f_i = F(u_i).C + F_{ext}(u_i) \quad (2)$$

where C is a constant factor that makes it possible to scale the height of the flame. So the flame profile will have a similar shape to this of the function F. F_{ext} is an additional forces function that adds a swinging effect. It can be a periodical function, noise or random function that allows some animation to be added and emphasizes the fact that this is a dynamic light.

4.3. Linear flames

The flame is built in a similar manner to the simple flame model, using the standard GLU NURBS interface. The control points matrix of the NURBS surface is composed of all the particles of the peripheral skeletons and the lowest and the highest particles of the lead skeletons. There is a line in the matrix for each peripheral skeleton. Its size is thus $N_p \times (M+2)$ where M is the number of particles in one skeleton. Each line of this matrix starts with the lowest particle of the nearest lead skeleton, then takes all particles of one peripheral skeleton and ends with the highest particle of the same lead skeleton. Denoting m the index of the particle in the skeleton, n the index of the skeleton, $L_{m,n}$ the particles of the lead skeletons and $P_{m,n}$ the particles of the peripheral skeletons, we can write the matrix A as shown in equation 1.

It can be noticed that peripheral skeletons $0, 1, N_p - 1$ reference the same lead skeleton because it is the nearest one for all of them. Peripheral skeletons $N_p/2-2, N_p/2-1, N_p/2$ also do so for the same reason. Figure 5 shows different flame profiles generated by different FDFs.

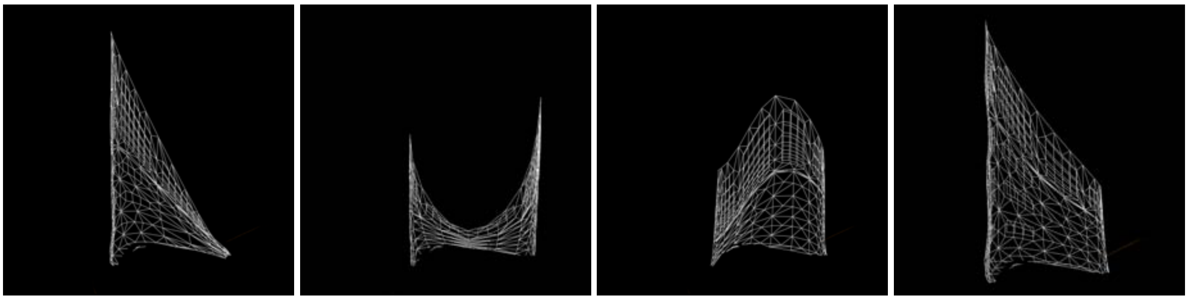
4.4. Complex flames

To build complex flames with NURBS surfaces, we assumed that any fire is a set of several independent flames. Indeed it seemed very difficult to represent a whole fire with a single animated NURBS surface. It's easier to achieve that with many smaller flames. Therefore any fire is defined by several linear flames.

An important point to understand is that we are not going to do anything on the geometry to merge the flames. That would be hard to do and above all very time-consuming. Thus any linear flame is built and rendered independently, and we will deal with their merging in a post-rendering phase described in the next section. Thus the wicks can be placed arbitrarily, and it will look even better if they are overlapping. Figure 6(a) shows a typical layout for the bottom of a torch and figure 6(b) for a campfire.

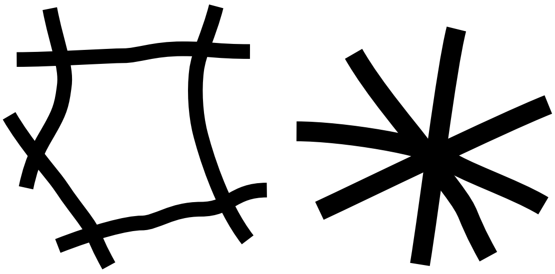
5. Rendering

We cannot rely only on 2D texture mapping to render multiple NURBS surfaces. As we said previously, the flames overlap, so the first thing to do is to deal with their transparency. Moreover we want a smooth appearance as far as possible, and most of all overlapping flames should visually merge together. Also, an important effect is the glow of the flame, because it is the only way to distinguish bright sources of light [NKON90].



(a) Linear (b) Bilinear (c) Gauss (d) Exponential

Figure 5: *Various linear flame profiles according to different FDFs with three lead skeletons*



(a) Torch model (b) Campfire model

Figure 6: *Top view of typical placements of the wicks*

5.1. True transparency

In the standard rendering pipeline of GPUs, we have to use alpha blending and sort the objects according to their depth from the back to the front if we want to see all the flames with transparency. Nevertheless this becomes quite tricky with our flames because they overlap and it is therefore often difficult to choose the first flame to be rendered.

That's why we implemented the method called *Depth Peeling*, which makes it possible to render objects with true transparency regardless of the order in n-passes. This was first realized using Virtual Pixel Maps [Mam89] and Dual Depth Buffers [Die96]. Later, graphics hardware functionalities allowed the technique [Eve02] to be accelerated. To display n transparency layers, the method uses n passes. The first pass retrieves the first depth layer with the usual depth test. The following passes extract the next layers using the depth from the previous pass to eliminate the previous depth layers.

The simplest way to explain its GPU implementation is to consider two depth buffers and two depth tests. The first depth test is the usual GL_LESS or GL_LEQUAL testing which keeps the current fragment if it is the nearest to the viewer. The second depth test rejects the active fragment if it is nearer than the depth of the pixel rendered at the previous pass. The first pass consists in rendering the scene with only the first test enabled, while the first depth buffer is saved in the second depth buffer. The next passes render the scene with both depth tests enabled. Each depth layer is rendered using additive blending.

With current graphics hardware, this can be done quite easily. We have used ARB_fragment_shadow extension to perform the second depth test quickly as well as the Frame Buffer Object (FBO) extension to avoid read-backs of the depth buffer, by attaching FBOs to depth textures. As we have now to render the flames many times for each frame, the cost of their rendering must be reduced. A straightforward way is to generate an OpenGL *display list* while rendering the NURBS surfaces in the first pass. Then this list can be used to render them in later passes.

However, our first results were slow because OpenGL kept doing the tessellation of the NURBS when using the display list. Indeed to really avoid tessellation, the NURBS surface must not be generated in GLU_NURBS_RENDERER mode. This has to be done in GLU_NURBS_TESSELLATOR mode, because that way we can define our own callback functions which will call the standard OpenGL functions to generate the vertices, normals and texture coordinates. That being done, the display list contains glVertex, glNormal, glTexCoord calls instead of the gluNurbsSurface calls. Thanks to that the frame rate drastically improved, for instance by a factor of four with four layers.

Last, the depth peeling algorithm has been slightly modified. Indeed, it assumes that all the scene is rendered in full transparency. But this is not exactly what we want, since we only want to apply it to the flames. Furthermore, rendering the whole scene would obviously slow down the process. Yet if the standard depth peeling algorithm is only applied to the flames, we will miss the occlusions of the other scene objects. When multiple flames are rendered, particular cases can even occur where an object in the scene is placed between two flames. That means that each pass of the depth peeling must perform a third depth test. We therefore pre-

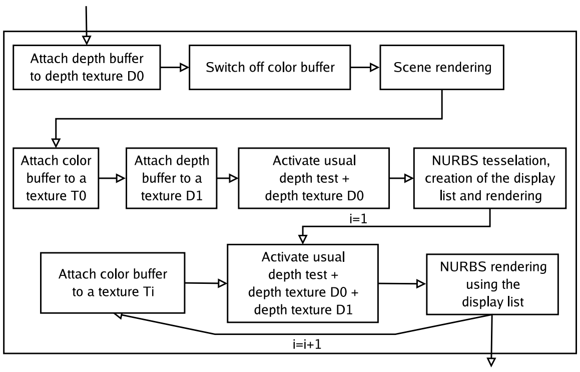


Figure 7: *Depth peeling process for flames*

render the scene without the flames in a FBO with color buffer writing disabled and depth buffer attached to a second depth texture. Thus a fragment is only displayed if it passes the two depth tests from the depth peeling and the depth test with this new depth texture. Figure 7 summarizes our implementation.

5.2. Glow

To render the glow of the flame, we implemented the technique described in [JO04]. It consists in a post-processing of the 2D rendering of the scene. Glowing objects are rendered separately. The frame buffer is rendered in a 2D texture at a lower resolution than the viewport. This texture is then filtered to produce a blur effect using a separable function, typically Gaussian. Thanks to this property, the blur can be performed in two passes. It is first performed in the x direction with n pixels and then in y with n pixels instead of $n \times n$ pixels in one pass. This makes it possible to do this on the GPU because this implies few texture lookups. Last the scene is rendered normally and the blurred texture is applied on top of it using additive alpha blending.

Our application of the algorithm is somehow similar. The following kernel is used to perform the blur in one direction :

$$K = \frac{e^{-x^2}}{\sigma^2} \tag{3}$$

where σ is equal to 1.5. These settings have proved to give good results. It is indeed important that the bandwidth is not too large, because it will produce too blurry a flame, and we would lose details in its texture.

However a blur at the same resolution as the viewport produced an interesting anti-aliasing effect, making the edges of the flame fairly smooth. We thus perform two different blurs. A first one is performed at the same resolution and a second at a lower resolution, generally four times smaller. The normal rendering of the flame is not used at all, we only render the two blurred ones. Note that to make the glow process efficient, it is crucial to render the flames only once, especially in this context because we use depth peeling which requires n passes for each rendering. That's why the rendering is stored in a texture which is the input of both blur processes. Figure 8 summarizes the glow process while figure 9 shows the whole rendering process.

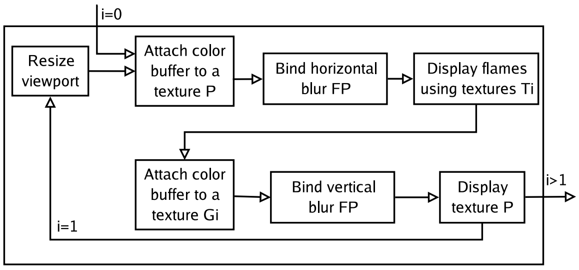


Figure 8: *Glow process for flames*

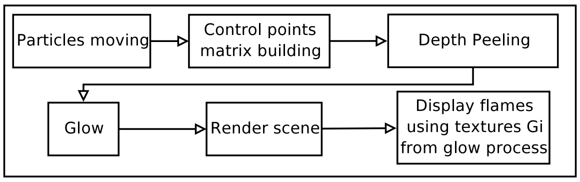


Figure 9: *Rendering process for flames*

6. Results

The same parameters as in [BLLRR06] are used for the solver. A resolution of $15 \times 15 \times 15$ for the grid is indeed enough and allows us to keep real-time frame rates. We can employ as many skeletons as we want for each flame. The larger they are, the more skeletons we need, but of course the more time consuming it is. Between ten and sixteen skeletons per flame is a good compromise. The skeletons themselves have a height of less than nine particles. The benchmarks were performed on a Pentium-M 2 Ghz with a NVIDIA GeForce Go 6600 graphics card and the screen resolution is 1024×768 pixels.

Figure 11 compares the rendering with and without the post-rendering processes. This highlights that the normal texture of the flame is not very bright. Indeed, as additive alpha blending is used in the depth peeling and the glow process, we must avoid saturation. Nonetheless, by managing this accumulation of intensities carefully, we can create the merging effect we were looking for. Although the assumption of a set of linear flames is not physically correct, we can see that it is finally visually convincing.

On the contrary to our previous paper, we did not consider the spatial distribution of the light. Indeed, it is harder to capture a photometric solid for the flames described

Figure 10: *A Roman draper shop illuminated by an oil lamp and a candle (67752 polygons, 2 independent flames, 25 fps)*

here than for candles. Thus we used a standard per-pixel lighting, using the spectral properties of a real flame. Figures 13 shows a campfire and figure 12 a torch in the baker shop of the forum. Both models produce 45 fps. Without the post-rendering processes, they reach 57 fps. The glow costs around 3 fps and the depth peeling with four layers 9 fps. Small flames can also benefit from this new rendering process (figure 10). The appearance of flames is however much better when animated, videos are available on our web site `http://www-lil.univ-littoral.fr/~bridault`.

7. Conclusion

We have presented a model that can handle complex flame rendering and animation in real time. The use of virtual wicks allows the user to put and merge flames onto any object easily. Far from here we have used our own software, but the NURBS surfaces used for modeling are available in standard graphics APIs. Thus our model can be integrated into any real-time application. The careful use of depth peeling and glow techniques helps to improve the realism of the flame rendering considerably, especially for complex flames like camp fires or torches that may appear in ancient illumination sources.

Future work will focus on different topics arising from flame simulation. Firstly we will study the problem of managing several flame-based sources. Indeed the fluid dynamic solver is still a bottleneck due to its high computational requirement. As soon as several independent sources are used, several independent solvers will have to be run, accordingly reducing the frame rate. We didn't deal with shadows rendering here yet, but as we stated in our previous paper, our model is compatible with common techniques. Smoke simulation involves finding a way to render its main curls and to consider participating media in order to simulate its propagation through the scene. Finally fire propagation could be simulated by adding a temporal dimension to our virtual wicks allowing them to be animated.

References

[BLLRR06] BRIDAULT-LOUCHEZ F., LEBLOND M., ROUSSELLE F., RENAUD C.: Enhanced illumination of reconstructed dynamic environments using a real-time flame model. In *Afrigraph '06: Proceedings of the 4th international conference on Computer graphics, virtual reality, visualisation and interaction in Africa* (New York, NY, USA, 2006), ACM Press, pp. 31–40.

[BPP01] BEAUDOIN P., PAQUET S., POULIN P.: Realistic and controllable fire simulation. In *Proceedings of Graphics Interface 2001* (2001), Watson B., Buchanan J. W., (Eds.), pp. 159–166.

[Cha02] CHALMERS A.: Very realistic graphics for visualising archaeological site reconstructions. In *Proceedings of the 18th spring conference on Computer graphics* (Avril 2002), ACM Press, pp. 43–48.

[DC01] DEVLIN K., CHALMERS A.: Realistic visualisation of the pompeii frescoes. In *Proceedings of the 1st international conference on Computer graphics, virtual reality and visualisation* (2001), ACM Press, pp. 43–48.

[Die96] DIEFENBACH P. J.: *Pipeline rendering: interaction and realism through hardware-based multi-pass rendering.* PhD thesis, Philadelphia, PA, USA, 1996.

[Eve02] EVERITT C.: Interactive order-independent transparency. tech. rep., 2002.

[HK03] HASINOFF S. W., KUTULAKOS K. N.: Photo-consistent 3d fire by flame-sheet decomposition. In *ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision* (Washington, DC, USA, 2003), IEEE Computer Society, p. 1184.

[IC03] IOANNIS R., CHALMERS A.: High fidelity lighting of knossos. In *VAST2003: 4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage* (2003), ACM SIGGRAPH Publications.

[IM04] IHRKE I., MAGNOR M.: Image-based tomographic reconstruction of flames. In *SCA '04: Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation* (New York, NY, USA, 2004), ACM Press, pp. 365–373.

[Ina90] INAKAGE M.: A simple model of flames. In *Proceedings of the eighth international conference of the Computer Graphics Society on CG International '90: computer graphics around the world* (Institute of Systems Science, Singapore, july 1990), pp. 71–81.

[JO04] JAMES G., O'RORKE J.: Real-time glow. *GPU Gems: Programming Techniques, Tips and Tricks for Real-Time Graphics* (2004), 343–361.

[LF02] LAMORLETTE A., FOSTER N.: Structural modeling of flames for a production environment. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (2002), ACM Press, pp. 729–735.

[Mam89] MAMMEN A.: Transparency and antialiasing algorithms implemented with the virtual pixel maps technique. *IEEE Comput. Graph. Appl. 9*, 4 (1989), 43–55.

[NFJ02] NGUYEN D. Q., FEDKIW R., JENSEN H. W.: Physically based modeling and animation of fire. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (2002), ACM Press, pp. 721–728.

[NKON90] NAKAMAE E., KANEDA K., OKAMOTO T., NISHITA T.: A lighting model aiming at drive simulators. In *SIGGRAPH '90: Proceedings of the 17th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1990), ACM Press, pp. 395–404.

[Per85] PERLIN K.: An image synthesizer. *SIGGRAPH Comput. Graph. 19*, 3 (1985), 287–296.

[Psz04] PSZCZOLKOWSKA D.: Visual model of fire. In *Eurographics 2004 Short Presentations* (Warsaw University of Technology Poland, 2004), 2004 E. A., (Ed.).

[Rac96] RACZKOWSKI J.: Visual simulation and animation of a laminar candle flame. In *International Conference On Image Processing and Computer Graphics* (1996).

[Ree83] REEVES W. T.: Particle systems : a technique for modeling a class of fuzzy objects. *ACM Trans. Graph. 2*, 2 (1983), 91–108.

[SF95] STAM J., FIUME E.: Depicting fire and other gaseous phenomena using diffusion processes. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques* (1995), ACM Press, pp. 129–136.

[Sta99] STAM J.: Stable fluids. In *Siggraph 1999, Computer Graphics Proceedings* (Los Angeles, 1999), Rockwood A., (Ed.), Addison Wesley Longman, pp. 121–128.

[Sta03] STAM J.: Real-time fluid dynamics for games. In *Proceedings of the Game Developer Conference* (2003).

[WLMK02] WEI X., LI W., MUELLER K., KAUFMAN A.: Simulating fire with texture splats, 2002.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Multi-Resolution Digital 3D Imaging System Applied to the Recording of Grotto Sites: the Case of the Grotta dei Cervi

J.-A. Beraldin[1], F. Blais[1], L. Cournoyer[1], M. Picard[1], D. Gamache[1]
V. Valzano[2], A. Bandiera[2]
M. Gorgoglione[3]

[1] IIT, National Research Council Canada, Ottawa, Ont, Canada, K1A OR6
[2] SIBA Coordination, University of Lecce, LE, Italy, 73100
[3] Superintendence of Archaeology of Apulia, Taranto, Italy

Abstract

The Grotta dei Cervi is a complex and fragile Neolithic cave where human presence left a large number of unique pictographs and petroglyphs. Detailed documentation necessitates recording it at different levels of details or spatial resolutions. A general approach would entail a combination of 3D data from different 3D sensors and information from different sources in order to meet set resolution targets. We used a prototype multi-resolution 3D laser imaging scanner that allowed acquiring the shape information of the three main chambers with a spatial resolution that improves with shorter standoffs. The system can record 3D data at a camera-to-object distance which ranges from 0.5 m to 10 m. At a standoff of 0.75 m, it provides a depth uncertainty of 0.08 mm and an optical lateral resolution of 0.2 mm on actual rock surfaces. This paper presents the project and the results obtained. The 10-day long visit into the Grotto generated more that 100 GB of 2D and 3D data that requires the development of new tools for modelling and managing the archive.

Categories and Subject Descriptors (according to ACMCCS): I.3.3 [Computer Graphics]: Picture/Image Generation - Digitizing and scanning

1. Introduction

1.1. High-resolution 2D and 3D information for multi-target applications

High-resolution 3D models of museum objects and heritage sites contain a wealth of information that can be examined and analyzed for a variety of conservation, research, and display applications [GBT*02]. For example, in the case of a site that must be closed or subjected to limited access for conservation reasons, an immersive 3D virtual reality theatre can be used to enable visitors to “virtually” visit the site. 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. In contrast to photographs, the actual geometric 3D position of each point on the surface of the model is available. Computer-based visual enhancement and analysis techniques can be applied to accomplish a precise “virtual restoration” that cannot readily be accomplished using traditional conservation techniques. For example, sections of paintings that have been removed from a grotto to a distant museum can be scanned and digitally reintegrated into the site 3D model for recontextualization. Enhancement techniques can be used to improve the legibility of faded images or inscriptions as well as to remove graffiti that has defaced the images. Finally, 3D models recorded before and after an actual conservation treatment, can serve as vital archival record for ongoing site monitoring and maintenance.

1.2. General approach to accurate 2D-3D site recording

The accurate recording of rock art sites, ancient crypts and grotto sites is a challenging task. The sites have either formed naturally or been carved from the surrounding rock and typically the walls, floors, and ceilings have an irregular surface shape and the paintings (pictographs) or carvings (petroglyphs) follow the contours of the rock surface over large areas. These features particularly the shape of the rock surface and speleothems (wall concretions, stalactites) are difficult to record with a high level of detail, measure, compare and display using conventional recording techniques such as survey methods, rectified photography, distance meters, etc. There is no single method for shape and appearance recording that works for all types of environment and at the same time is fully automated and satisfies the requirements of the application. A general approach combines models created from multiple images, single images, range sensors, known shapes, CAD drawings, existing maps, and survey data [EBG*05]. The main objective is to minimize the impact of measurement uncertainties, augment the amount of information available (spatial resolution), and reduce costs and time spent on those sites.

Figure 1: *Grotta dei Cervi, Neolithic cave, Italy (40°04'47"N, 18°29'02"E). Photographs showing pictographs (Guano, Okra), petroglyphs and speleothems.*

1.3. Neolithic Cave: Grotta dei Cervi, Italy

The *Grotta dei Cervi* project started in February 2004, and it aims at recording the shape and appearance of that cave and to push 3D technology to higher levels. The site is composed of three main corridors decorated with Neolithic pictographs made of red okra and bat guano and petroglyphs (see Fig. 1) [Gra02]. The Grotto discovered in 1970 by local speleologists is located in South-eastern Italy at Porto Badisco (40°04'47"N, 18°29'02"E), Otranto, LE. The main entrance is situated at 26 m above sea level and the largest depth is about 26 m. The Grotto contains a rich stygobitic fauna. The temperature is fairly constant at 18°C and the RH hovers between 98% and 100% [LGS*00]. The main motivation for this project comes from the fact that the cave is closed to the public and only a limited number of experts are allowed in every year. This measure is necessary in order to preserve the delicate environmental balance inside the site. Consequently, a detailed 3D model draped with colour information would allow for increased information on the site through detailed studies and virtual visits without traumatic consequences to it. Work on the central corridor measuring about 300 m long is underway (see Fig. 2). This corridor has passage ways barely allowing an adult in and some chambers have a maximum cross section of about 8 m wide × 5 m high.

1.4. Organization of paper

The objective of this paper is to review current literature on rock art 3D recording, and, present the project, an overview of the multi-resolution 3D imaging system used and the results obtained up to now. This research work is in progress because the 10-day long visit to the Grotto generated more that 100 GB of high resolution 2D and 3D data that will require the development of new tools for modelling and managing the archive. Section 2 summarizes a number of projects aimed at documenting in 3D rock art sites. Section 3 presents details about the planning of the project. Section 4 discusses the characteristics of the prototype laser scanner and Section 5 will describe the preliminary results. This paper ends with some concluding remarks.

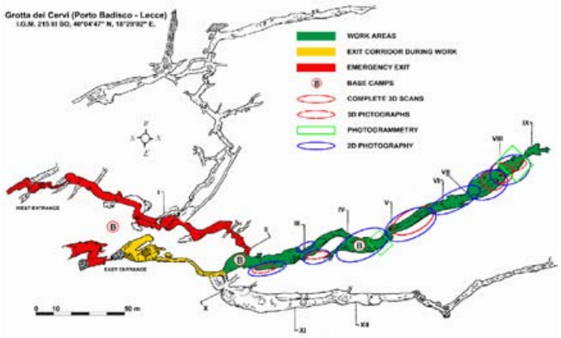


Figure 2: *Map of the cave and illustration of the areas completed (modified from [Gra02]).*

2. Some projects on 3D recording of Rock Art sites

These examples will show the challenges in preserving record rock art sites and the increased interest in using laser scanners amongst other technique to acquire dense high resolution 3D information.

2.1. Caves and Crypts

As early as 1994, laser scanning has been used for cave recordings [Men02]. Electricité de France and Mensi undertook the modelling of the Cosquer cave that was discovered in 1985 by a diver near Marseille (France). The cave walls are decorated by paintings that are between 19000 and 27000 years old. Access to the cave is treacherous: the entry is 37 meters below the water level with a 175-meter long passage tunnel. This cave is now protected; the access is closed. The triangulation-based laser scanner was transported in a water proof case into the cave. The computer was positioned on the coast and connected directly to the scanner via a 300-meter cable. The final model was created from 28 different scanner positions, 128 images from the on-board video colour camera and took 67 work hours to create a model with 4.7 millions 3D points – resolution of 30 mm (XY) and 1 mm at 5 m (Z). In 1996, in collaboration with the Israel Antiquities Authority, NRC Canada undertook a project to demonstrate the application of a portable triangulation-based laser scanner for conservation documentation of the Arcosolia Room of the Tomb of St. James and a number of other sites in Israel [GBT*02]. The Arcosolia Room of St.

James Tomb, which measures approximately 2 m x 2 m x 1.8 m, has been carved in the rock, and its interior surfaces are rough and irregular. The objective was to digitize the entire interior of the Tomb to prepare an archival record for conservation documentation. The entire interior of the Tomb was recorded at 10000 3D points per second with a lateral resolution of 2 mm (X,Y) and a depth uncertainty (Z) of 0.3 mm in one half day of on-site recording time. Subsequently, approximately 4 to 5 days were required off-site to prepare the 3D digital model.

In 2001, the SIBA Coordination at the University of Lecce, Italy in collaboration with NRC Canada applied a technique, which combined photogrammetry, 2D digital photography and 3D scanning to prepare a photo-realistic 3D model of the Byzantine Crypt of Santa Cristina in Carpignano, Italy for conservation documentation and visual communications [BPE*05]. The crypt excavated around the 9th century C.E., measures about 16.5 m x 10 m x 2.5 m and has a number of well-preserved frescoes on the walls. A photogrammetric technique was used for the main entrances and a commercial triangulation-based laser range scanner was used to provide 3D points at a rate of 100 points per second for the interior. Texture was acquired with a 6 mega-pixel digital camera. A total of 92 hours were spent in the Crypt to acquire about 12.8 million 3D points. The 3D spatial resolution on the walls is about 5 mm and on the ceiling and floor, 15 mm; texture resolution is 1 mm on the walls. The range uncertainty was estimated at 0.8 mm. The 3D model was created over a period of one month. A CDROM and DVD about the site along with a movie showing a virtual fly through of the Byzantine Crypt were prepared.

The Altamira cave located in Spain was a popular stop for tourists up to 1970. It was then closed because of the high number of visitors that resulted in an increase in temperature and humidity which is known to cause the prehistoric cave paintings to flake off the walls. The cave authorities decided to reconstruct the whole cave true to size so that the public could view the prehistoric paintings without risking damage to the original. A close range triangulation-based laser scanner recorded three-dimensional images of more than 2600 m² of painted walls [Min01]. Negative shapes were cut in foam and silicone moulds were made. The silicone moulds were painted by hand with natural pigments. The completed physical cave model is cooled down to 12°C to give a genuine cave experience. No indications of spatial resolution were found. Others have also been interested to laser scanning mapping in other caves around the world. In [KKC01], the authors discuss a multi sensory 3D scanning method where they achieve a fully integrated 3D model with texture that has roughly a 2 mm mesh density. In [GVI*05] the authors describe a solution for the remote fruition of a cave site based on a mobile robot. They aim at increasing both off-line and on-line experience of such sites with a robot that moves in the cave and collects both colour and 3D structures. They performed their experiment in the “Grotta dei Cervi” (“Stag’s Cave”) zone III. At the moment, spatial resolution and accuracy figures have not been released.

2.2. Pictographs and petroglyphs

In [EFP04], the authors present an approach to create detailed and realistic 3D models of Aboriginal pictographs of the Baiame cave in New South Wales, Australia. They used a combination of time of flight laser scanning, bundle adjustment, and surveying. The technique achieves the texture mapping without extracting common points between the texture images and the 3D geometric model. A total station was used to “tie” the data together by measuring the coordinates of points which were discernable in the data. The standard deviations of the computed 3D coordinates were 13 mm (X), 9 mm (Y), and 11 mm (Z). In [CF05], according Chandler and Fryer, archaeologists, conservators and rock-site managers need record sites using simple and cost effective methods. They developed a methodology that enables a non-expert to acquire images suitable for photogrammetric measurement using a 3-megapixel digital camera and a scale bar. No laser scanner is necessary or desirable. At the moment, they still require expensive professional grade software and an external self-calibrating bundle adjustment to generate both accurate and dense DEMs/orthophotographs. They conducted fieldwork experiments in New South Wales, Australia where both petroglyphs and pictographs were recorded. The petroglyphs measure less than two meters in extent and the stereo imagery was captured by raising the digital camera 1 to 2 meters above the surface. A second 6-megapixel camera was used to assess the accuracy of the cheaper camera. In sharp contrast, the authors in [WSVW04] take a route based on a mid range laser scanner and GPS surveying instrument to record petroglyphs in Little Lake, CA, USA. At the moment, the topic of modelling petroglyphs by either laser scanning or photogrammetry is being revisited by many researchers around the world. The main reasons are that laser scanning is still seen as expensive, difficult to transport into the field and requires some expertise to use successfully during data capture and photogrammetry gives rather inaccurate and sparse data on featureless surfaces.

2.3. Illuminating rock art

In [Cha02], the author oriented his work to the creation of highly realistic rendered images which are based on as much real-world data as is obtainable. For the author, much research has been undertaken into accurately modelling archaeological sites and reconstructing incomplete structures. Unfortunately, the modelling of the illumination used by standard software packages for rendering tend to be based on sources that simulate daylight, and artificial sources, rather than original flame-based sources like oil lamp or candle light. It is essential that these be modelled accurately when recreating these ancient conditions. The site used to test the approach is a section of the cave at Cap Blanc, France which dates back to the Upper Palaeolithic era. The authors in [MYKI05] simulated how natural light might have illuminated the interior walls of the Fugoppe Cave, located in northern Japan which they obtained from their proposed modelling system based on short and mid range laser scanners. They examined the possibility that ancient painters and sculptors worked under natural light as opposed to what archaeologists believe that they worked using artificial lights such as torches.

3. Grotta dei Cervi: Practical considerations

A reconnaissance visit to the grotto in 2004 allowed the team to plan the activities. The main concerns were the technical difficulties the team might encounter and the determination of the required spatial resolutions. We had strict guidelines to abide to before entering the Grotto, a few are listed here

- emergency exits, equipment and first aid stations had to be determined in advance,
- drop cloths and carpet made of a resilient mesh to be used for equipment and people,
- light weight wires for electricity and communication located in fixed positions,
- constantly monitor temperature during work,
- two speleologists present in the grotto at all time,
- no modifications allowed to the site to fit the equipment or people.

The spatial resolution of a model depends on the level of details sought by the intended application of the 3D model but also by the equipment available and by practical and other logistic constraints. With reference to Figure 2, a list of interesting areas was considered for recording:

- Low and high-resolution 3D scans in zones III and VIII,
- Complete scanning of zones V and VII,
- Very-high-resolution 3D scanning of the Shaman of section VIII (2000 x 2000 pixels image),
- Lower and medium resolution 3D scans of other sections of the caves and passages that have less historical significance but important to create a more realistic 3D model of the site; some low-resolution scans of some floors were also done,
- High-resolution texture photographs using flash lamps according to work above,

In a few months during 2004, we put together the financing, the equipment, and complete team for the February 2005 visit. The NRC team was composed of five individuals directly involved in the project and one programmer that stayed on-line in Canada. From the SIBA-University of Lecce, five individuals participated directly to the recording inside the Grotto and one programmer that stayed on-line outside. Two experienced speleologists and two archaeologists accompanied and took an active role during the recording. The project went according to plan with minimum modifications from the original schedule even when considering the complexity and difficulty of this very challenging environment: dust, humidity, and size, etc. To limit the quantity of equipment inside the Grotto and to adapt quickly to the unpredictability of the irregularly shaded walls, we used our extensively modified prototype high resolution 3D laser imaging scanner. It acted like a three-in-one laser scanner. Section 4 will describe this prototype laser scanner. For texture acquisition, one digital camera (Kodak-14n®) with two lenses (24 mm) and four 500 watts•sec flashes were used. A low noise power generator, one UPS, two parallel electrical power lines of 300 m long: one for the 3D laser scanner, and the second for the flashes. One Ethernet link insured constant cave to surface communication, backup and data transfer between the computers. Some redundancy proved necessary and provided very smooth operation that otherwise may have jeopardized the whole project. The UPS placed outside the Grotto provided constant uniform voltage regulation needed for the proper operation of the laser scanner which was completely isolated from the four flashes.

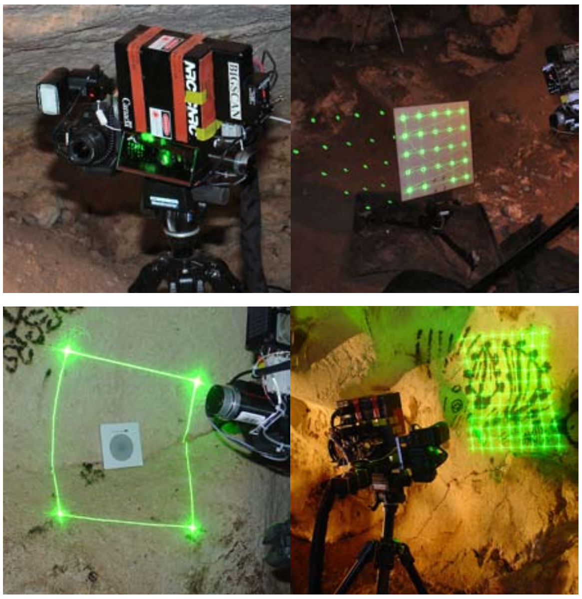


Figure 3: *Photographs showing the 3D laser scanner, calibration and resolution objects used to constantly verify and track cameras performance, virtual calibration grid for 2D camera.*

4. Modelling techniques used for the project

In general, the selection of a particular 2D-3D solution must consider that the geometric fidelity depends mainly on the accuracy and data noise (measurement uncertainty-Z) and the level of surface details that can be sensed by the scanner/camera (spatial resolution-X-Y).

4.1. On the physics for optimal acquisition of shape and colour texture information

Without an understanding of the physics behind recording, most 3D acquisition and modelling methods will impose serious limitations on the final quality of the models and potential virtual interactions with them. Simply applying the typical processing pipeline used for 3D modelling will result in many cases in crude renderings of the 3D model caused by excessive filtering or polygonal compression or blurring because of poor image resolution. According to [BB06], the main reasons behind these limitations were until recently imputed to software that could not handle large models (> 50 million polygons). But today, displaying these models is being solved and one may now argue that increasing the number of recorded points or buying higher resolution imagers will solve the question of resolution. This simple answer is naive, it is hiding a more important issue which is the initial quality of the raw images being acquired; it is completely useless to over-sample if the 3D and textures are blurred to start with.

The authors in [BPE*05] present an overview of the basic theory about 3D sensing to help the reader decide on which technology to use. Blais in [Blais04] reviews many scanners on the market (prior to 2004). Interesting enough, for range between 5 m and 10 m, there are very few laser scanners available commercially that are adequate for very high resolution visualization application. This span of distances represents a transition between triangulation and time of flight systems.

Figure 4: *Team member and scanner at work.*

Figure 5: *Preliminary result of a very high-resolution scan (X-Y: 0.2 mm spacing) performed over the Shaman in Zone VIII. The 3D image is shown without colour and using side shading from an artificial light source.*

Obtaining colour texture using RGB lasers, collinear colour cameras, and perspective projection methods (texture mapping) has successfully been demonstrated [ABC04, BB06, BPE*05, GBT*02] and some of these techniques are currently being commercialized. Textures acquired by separate digital camera are registered with the geometric model using common points between the 2D image and the 3D model. This must be done for each image unless the camera is fixed to the scanner, then it may only be done once. In the latter situation, factory calibration is acceptable only if the two systems are rigidly mounted and do not require adjustments. This is valid for specific configurations such as systems with a fixed focus camera lens, usually focused at infinity for convenience e.g. in long-range laser scanners. However, for close-range measurements, this is not acceptable. The high resolution colour camera must constantly be optimized by either re-focusing the lens and/or adjusting the converging angle to better match the field-of-view of the colour camera and 3D scanner. Increasing the number of pixels on the 2D camera certainly helps but is by far not sufficient because of the physical limitations imposed by the camera depth of field and field of view. The complex interrelations between optimal field of view, numeral apertures, focal length of the lens and physical parameters of the detector such as the number of pixels and size were investigated by [BB06]. In that reference, charts and equations are provided as guidelines to facilitate acquiring high resolution 3D and colour images. It is shown that compromises are mostly imposed by the laws of physics rather than engineering.

Figure 6: *Same area but after mapping the colour photograph (Canon-20D®) onto the 3D image.*

4.2. NRC's 3D technology

The scanner used for this project is known as "Big Scan" laser scanner. It is a research prototype system currently under development for high-resolution 3D digitization of large structures. The system allows 3D recordings at a camera to object distance which ranges from 0.5 m (camera standoff) to 10 m. At a standoff of 0.75 m, it provides a resolution of 0.08 mm on cooperative surfaces. A space qualified version of this scanner was built by Neptec design Group for inspection tasks on in-flight space shuttle missions. The use of this technology both for large-scale rock face analysis and for small-scale, detailed analysis and geo-material classification has been studied. The 3D NRC Laser Scanner system uses a green (532 nm wavelength) laser to acquire high-resolution 3D images of the surface of the object (Figure 4), a red laser marker was used to help manage the acquisition and the overlap between the images. 3D lateral resolution varies depending on the distance of measurement and the point density of the 3D images:

- 0.2 mm @ 0.75 m in the very high-resolution mode (for a 2048 x 2048 pixels images, Fig. 5-6),
- 0.4 mm @ 0.75 m for the high-resolution mode (1024 x 1024 pixels),
- 0.75 to 1.0 mm @ 1.4 to 2.0 m for the medium-resolution mode (1024 x 1024 pixels),
- 2 mm @ 2 m in the low-resolution mode (512 x 512 pixels).

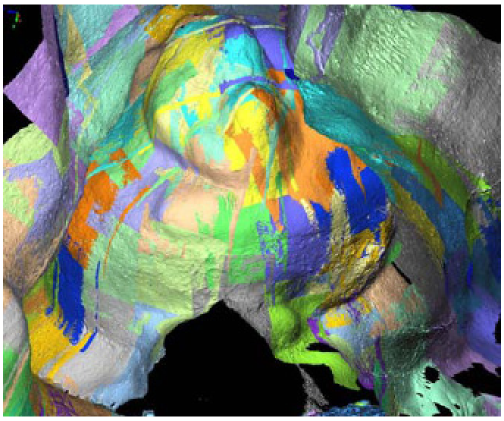


Figure 7: *View from one end of the inside of zone VII looking at the entrance to zone VIII. This view of zone VII is shown with independent colours for each scan. It illustrates the overlapping areas between images.*

5. Preliminary results

We returned inside the Grotto in February 2005. Here, we summarize the main results obtained so far.

<u>3D scanning</u>

The 3D model is created by acquiring a mosaic of 3D images sections and by stitching these images by software. Overlapping areas between the 3D images is needed in order to properly align them together. The operation illustrated in Figure 7 was performed in the grotto to verify the quality of the image and pre-aligns the 3D images to avoid missing data (using IMAlign® from the Polyworks® Suite). Figure 8 shows the same area using synthetic shading (the colour information has been removed). An 8-Megapixel camera (Canan20D®) is mounted on the laser scanner (Fig.3). The camera intrinsic and extrinsic parameters are calibrated on the fly using a virtual 3D grid generated by the laser camera itself (Fig.3).

<u>2D texture photography</u>

As shown on Figure 9, 2D photographs of key critical sections of the grotto were acquired using a 14-Megapixel digital camera (24 mm lens focal length) at two predefined pre-calibrated positions: 2.0 m for close-ups and infinity for global image modelling. These images will be used to:

- Provide very-high resolution (4500 × 3000 pixels) colour texture mapping on the 3D models for each section of the “Grotta”.
- Good consistency in the reference colour and better uniform illumination.
- Accurate colour texture of sections of the cave that were not accessible by the laser scanner.

At the moment, these high-resolution photographs must be manually mapped on a 3D object using software packages such as TexCapture®.

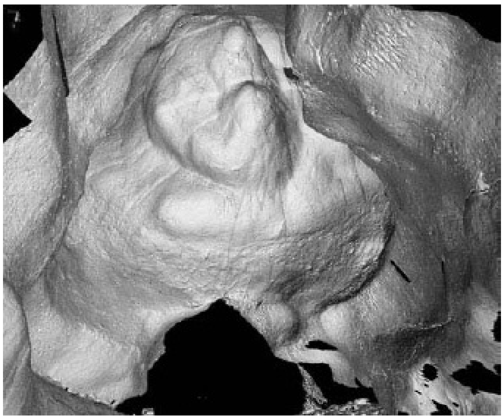


Figure 8: *Same portion of zone VII scanned at medium and low resolution and shown with artificial shading.*

Figure 9: *Team working on the photographs for the texture of zone VIII, camera and flashes are shown.*

Figure 10: *Preliminary result of a High-Resolution scan made of a number of 3D images (X-Y: 0.5 mm spacing) performed in zone V.*

Figure 10 shows a preliminary result obtained and illustrating the complexity of the 3D shape of the Grotto. It is important to mention that the colours were not corrected and that these images are only pre-corrected and pre-aligned. A total of 35 GBytes for 716-3D images with photographs and 65 GBytes for 1786-high-resolution colour texture photographs were acquired.

Table 1: *Project "Grotta Dei Cervi" in numbers.*

Number of rooms scanned	4
Smallest/largest objects scanned	0.5 m to 6 m
Smallest spatial resolution	0.2 mm for geometry, 0.2 mm for colour
Number of 3D images acquired	716 X-Y-Z
Number of 3D points acquired	630 million points
Number of 2D images acquired	3500 images
Total amount of data captured	100 gigabytes
Scanner weight	3 kg
Length of electrical cables Length of Ethernet cables	600 m 300 m
Deepest work level underground	26 m
Narrowest passage	0.6 m x 0.6 m

Accurate registration of the 3D model, colour texture projection and compensation are planned for a phase II of the project. This very large amount of high-resolution 2D and 3D image data opens the door to providing a 3D model of the caves of unmatched resolution, never obtained before in any 3D model of a grotto site. This will be a major scientific milestone in modelling large and complex 3D environments. Our major challenge is associated with the size and resolution of the 3D images which causes computer crashes and excessive processing time. As a rule of thumb, the memory requirement increases linearly as the complexity of the model augment and in a quadratic manner for the processing. We are pushing the image resolution by almost an order of magnitude compared to previous work which implies that even the limits imposed by the computer operating system, for example 1.7GBytes for Windows, is becoming a major problem.

6. Conclusion

The potential of modelling as-built heritage sites opens-up promising applications. As demonstrated with the *Grotta dei Cervi* project and many other projects worldwide, a high degree of realism can be attained by those techniques. Three-dimensional models of ancient rock art sites, ancient crypts and grotto provide an important new level of documentation, which can be used for a variety of conservation, research and display applications. Perhaps the most important is that 3D VR theatre displays of accurate virtualized models of those sites can be used to enable very realistic virtual visits to the sites in lieu of actual site visits, which endanger the site itself. A second very important application is the use of the data to reliably monitor the condition and stability of the site. The problem we addressed and the approach we proposed in this paper are aimed at the effective use of 3D modelling to enhance the understanding of a heritage site that needs to be preserved and shown to more people in order to raise awareness and understanding of rock art sites that are fragile, inaccessible and usually located in remote areas. The work will continue and we expect to do more research to create tools that will handle the models that will be created from the 100 GB of 2D and 3D data generated in the course of the 10-day visit to the grotto. Those models will be used for public outreach purposes and experts will have access to the data through collaborative projects. More research work is required to speed up the process of acquisition and modelling. These operations still require a larger amount of time. But one needs to understand that the current level of recording efforts and cost are worth spending compared to risking the forever loss of important historical sites to vandalism, natural disasters or wars.

7. Acknowledgements

The 3D recording of the Grotta dei Cervi was realized within I18 of the "Piano Coordinato delle Università di Catania e Lecce" co-financed by the European Union (FESR, PON Ricerca 2000-2006). We would like to thank L.G. Dicaire, I. Cancelliere, M. Caputo, G. Ciccarese, A. Malcangi, S. Martiradonna, F. Melcarne, S. Nuccio, P. Pulli and A. Toma through their great dedication made important contributions to this project. Financial support was also made by the town Council of Otranto (LE), the CEDAD and the Department of Beni Culturali of the University of Lecce, the Museo Provinciale of Lecce and the CASPUR of Rome.

References

[ABC04] ANDREETTO M., BRUSCO N., CORTELAZZO G.M.: Artifact removal in 3D textures. In *Proc. CVMP '04* (2004), V. 1, 111-117.

[BB06] BLAIS F., BERALDIN J.-A.: Recent developments in 3D multi-modal laser imagng applied to cultural heritage. *Machine Vision and Applications*, in press.

[BPE*05] BERALDIN J.-A., PICARD M., EL-HAKIM S.F., ET AL: Virtual reconstruction of heritage sites: opportunities & challenges created by 3D technologies. In *Proc. Int. Workshop on Rec., Modelling and Vis .of Cultural Heritage*, Ascona 2005, 141-156.

[Blais04] BLAIS F.: Review of 20 years of range sensor development, J. Elec. Ima., 13(1), Jan. 2004, 232-240.

[CF05] CHANDLER J.H., FRYER J.G.: Recording Aboriginal Rock Art Using Cheap Digital Cameras and Digital Photogrammetry. In *Proc. CIPA '05* (2005), 193-198.

[Cha02] CHALMERS A.: Very realistic graphics for visualising archaeological site reconstructions. In *Proc. 18th spring conf. on Comp. graphics '02* (2002), 43-48.

[EBG*05] EL-HAKIM S.F., BERALDIN J.-A., GONZO L., ET AL.: A hierarchical 3D reconstruction approach for documenting complex heritage sites. In *Proc. CIPA '05* (2005), 790-795.

[EFP04] EL-HAKIM S.F., FRYER J., PICARD M.: Modeling and Visualization of Aboriginal Rock Art in the Baiame Cave. In *Proc. ISPRS '04* (2004), 990-995.

[GBT*02] GODIN G., BERALDIN J.-A., TAYLOR J, ET AL: Active Optical 3D Imaging for Heritage Applications. In *Proc. IEEE-CGA*, 2002, Vol. 22, No. 5, 24-36.

[Gra02] GRAZIOSI, P.: *Le pitture preistoriche della grotta di Porto Badisco*. Rist. anast. Istituto Italiano di Preistoria e Protostoria, 2002.

[GVI*05] GRAMEGNA T., VENTURINO L., IANIGRO M., ET AL.: Pre-Historical Cave Fruition through Robotic Inspection. In *Proc. IEEE Int. Conf. on Robotics and Auto.'05* (2005), 3187-3192.

[KKC01] KANAYA I., KADOBAYASHI R., CHIHARA K.: Three-Dimensional Modelling of Shofukuji Burial Chamber. In *Proc. VSMM '01* (2001), 113-120.

[LGS*00] LAIZ L., GROTH I., SCHUMANN P., ZEZZA F., ET AL.: Microbiology of the stalactites from Grotta dei Cervi, Porto Badisco, *Intern. Microb. 3* (2000), 25-30.

[Men02] MENSI: Cosquer Cave in Marseille. Cave Scanning Report. In CD-ROM *Cultural & Multimedia Presentation*, © Mensi 2000.

[Min01] MINOLTA EUROPE GMBH. *3D Digitizing. Altamira II: high-tech cave reconstruction*. 2001.

[MYKI05] MASUDA T., YAMADA Y, KUCHITSU N., IKEUCHI K.: Sunlight Illumination Simulation for Archaeological Investigation - Case Study of the Fugoppe Cave. In: *Proc. CREST '05*, 65-72.

[WSVW04] WASKLEWICZ T., STALEY D., VOLKER H., WHITLEY D.: Terrestrial 3d Laser Scanning: A New Method For Recording Rock Art.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Terrestrial laser scanner and high-resolution camera registration through single image-based modeling

D.G. Aguilera and J. G. Lahoz

High Polytechnic School, Avila, University of Salamanca, Spain

Abstract

This paper deals with an important topic: the automatic co-registration of terrestrial laser scanner data and high-resolution digital images. Our approach harnesses the power of a single image-based modeling method developed focusing on obtaining a spatial dimensional analysis from a single image. Particularly, the problem of image registration is solved automatically through a camera calibration method which takes 2D and 3D points correspondences as input data based on a search of spatial invariants: two distances and one angle.

1. Introduction

1.1 Context

The technological development in the last years has made possible the improvement of systems for geometry and colour object's measurements. From a sensorial point of view, active and passive techniques based on terrestrial laser scanners and high-resolution cameras have monopolized this leadership respectively. Thus, the demand of 3D models for objects documentation and visualization has drastically increased. 3D modeling of close-range objects is required in manifold applications, like cultural heritage, industry, cartography, architecture, archaeology, civil engineering, medicine, and last but not least, tourism and can be accomplished with traditional image-based modeling approaches or with scanning instruments.

Particularly, the image-based modeling pipeline constitutes a very portable and low-cost technique which consists on the 3D reconstruction of objects from one or more images. In this sense, several assumptions have to be solved: from camera self-calibration and image point measurements, to 3D points cloud generation, surface extraction and texturing. In this way, image-based modeling is a technique that has undergone a big growth in the last years. This promising evolution could be portrayed by the following issues:

- New technological neighbors and new relations among these: Photogrammetry, Image Processing, Computer Graphics, Computer Vision, etc.

- New algorithms and methods have emerged in order to achieve automatization and provide new products.

On the other hand, terrestrial laser scanning methods allow to recover directly 3D measurements of the scanned scene in a few seconds, providing a high level of geometric details together with a good metric accuracy. However, up to now the 3D reconstruction of precise and reliable large objects and scenes from unorganized points clouds derived from laser scanner is a very hard problem, not completely solved and problematic in case of incomplete, noisy and sparse data. As a result, nowadays none scanner can fulfill all demands in 3D modelization projects. Although the measuring process is very fast and simple, users should be well aware that, in addition to an appropriate software, time and patience are needed to get a final result in the form of a CAD drawing or a surface representation based on a topological triangulated mesh. The high complexity of 3D modelization requires a flexible multi-input and multi-output approach able to support the information arising from different sensors/techniques and to provide different levels of information to users with different requirements [FIM*05]. In this way, the key pass through taking advance of the opportunities open by the new communication and information technologies, as well as exploit the synergies with other disciplines in order to establish specific tools.

To reinforce this need, next table (Table 1) illustrates a comparison based on the most important features with relation to laser scanning and image-based modeling methods.

Laser Scanning	Image-based modeling
↓ Inaccurate lines and joints	↑ Accurate lines and joints
↓ Poor colour information	↑Good colour information
↑ Prompt and accurate metric information	↓ Hard-working and slow metric information
↑ Excellent technique for the description of complex and irregular surfaces	↓ Time-consuming technique for the description of complex and irregular surfaces
↓ High-cost technique	↑ Low-cost technique
↓ The 3D model is an entity disorganized and without topology	↑The 3D model is an entity organized and with topology
↑ Light is not required to work	↓ Light is required to work

Table 1: *Comparison of features: Laser scanning vs. Image-based modeling.*

The question, which technique is 'better' than the other, cannot be answered across the board. As we can see (Table 1), each technique owns its advantages at different working fields. In many cases, a combination of both techniques might be a useful solution.

1.2 Related work

In this integration of techniques, where a 3D scanner is used to acquire precise geometry and a digital camera captures appearance information, the 3D model and images must be registered together in order to connect geometry and texture information.

This problem of image to model registration is closely related to the problem of camera calibration, which finds a mapping between the 3D world (object space) and a 2D image. This mapping is characterized by a rigid transformation and a camera model, also referred to as the camera's extrinsic and intrinsic parameters. This rigid body transformation takes 3D points from object space to 2D points in the camera's reference frame, and the camera model describes how these are projected onto the image plane.

The camera calibration problem is solved by matching features in the 3D model with features in the image. These features are usually points, lines or special designed objects that are placed in the scene. The matching process can be automatic or user driven, and the number of feature pairs required will depend on whether we are solving for the intrinsic, extrinsic or both parameters sets.

In the context of image registration for 3D modeling using dense laser scanner data, several approaches have been developed up to now.

A pre-calibration of camera which allows to integrate geometry and texture avoiding any user post-processing used for the Digital Michelangelo project [LPC*00], or the approach described by [RCM*99] where the image to model registration is done manually by a user who selects corresponding pairs of points. Both approaches are applied in a context of small object modeling.

In search of an automatic method, [LHS01] develop an image registration approach based on silhouette matching, where the contour of a rendered version of the object is matched against the silhouette of the object in the image. No user intervention is required, but their method is limited to cases where a single image completely captures the object.

In other scale of methods applied to large distances, dealing with outdoor scenes and based on locating invariant image features, [MNP*99] suggest correlating edges common to the color image and the range map's intensity component. [Els98] aligns images by matching up the corners of planar surfaces. More recently, [SA01] present an automatic method for image to model registration of urban scenes, where 3D lines are extracted from the point clouds of buildings and matched against edges extracted from the images. [INN*03] in their Great Buddha work, use reflectance edges obtained form the 3D points and match them against edges in the image to obtain the camera position. Finally, [ATS*03] present a novel method for 2D to 3D texture mapping using shadows as cues. They pose registration of 2D images with the 3D model as an optimization problem that uses knowledge of the Sun's position to estimate shadows in a scene, and use the shadows produced as a cue to refine the registration parameters.

In a similar context, our approach harnesses the power of a single image-based modeling method developed in [Agu05] focusing on obtaining a spatial dimensional analysis from a single image. Particularly, the problem of image registration is solved automatically through Tsai calibration algorithm [Tsa89] which takes 2D and 3D points correspondences as input data based on a search of spatial invariants: two distances and one angle.

2. Multi-sensor description

The Trimble GS200 laser scanner (Figure 1) was employed for the scanning process. This scanning system is provided with a rotating head and two inner high speed rotating mirrors that allow to acquire a scene with a large enough field of view, i.e. 360º H x 60º V, reducing the need of using lots of scan stations. The sensor accuracy is below 1.5mm at 50m of distance with a beam diameter of 3mm. Furthermore, the laser allows to acquire reflected beam intensity and RGB colours.

A high-resolution camera, Nikon D70 (Figure 1), was used to overcome the poor colour information obtained from terrestrial laser scanner.

Figure 1: *Trimble GS200 laser scanner and digital camera, Nikon D70.*

3. Multi-sensor registration through single image-based modeling

A hierarchical process supported by single image-based modeling has been developed in order to register high-resolution images with laser scanner models. Nevertheless, before a 3D model can be texture mapped with a colour image, the transformation that aligns the two datasets must be estimated, which is not an easy task. The registration process is difficult to automate because image and laser points cloud are dataset which arise from sensors with different features: from its own intrinsic characteristics to features like its resolution and field of view.

The main contribution of this paper is the adaptation of a single image-based modeling approach in order to obtain geometrical constraints and a spatial dimensional analysis, which allow performing image to laser model registration automatically.
Our approach exploits vanishing points geometry inherent in oblique images as well as some geometrical constraints typical in architectural scenes. Particularly, four main steps are resolved sequentially: the first step involves an image analysis procedure based on recognition, extraction and labeling of features (special targets and vanishing lines); the second step involves the estimation of camera calibration exploiting vanishing points geometry; the third step carries out a dimensional analysis derived from a single image. This step uses the estimation of camera calibration, as well as some geometrical constraints used in single image-based modeling. Finally, the fourth step involves a search of correspondences between both dataset (3D points cloud and 2D image points) based on analyzing spatial invariants between special targets. This last step provides image to model registration together with a camera calibration tuning.

Nevertheless, this approach is only successful in a given domain where the following assumptions have to be considered:

a) The method is applicable in scenes with strong geometric contents such as architectural scenes.

b) The images acquired by digital camera have to be oblique with at least two vanishing points.

c) Special planar targets (Figure 4) are used as landmarks and have to be fixed to the building facades.

d) In order to have a primary camera pose estimation from a single view, user must know some priori information about the object (i.e. a distance) which performs as the reference frame.

The following scheme (Figure 2) aims to illustrate the methodology that we have developed in order to obtain multi-sensor registration through single image-based modeling automatically.

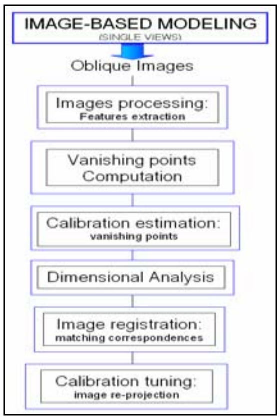


Figure 2: ***Multi-sensor registration through single image-based modeling.***

3.1 Image processing: Features extraction

A hybrid image processing step which integrates lines (vanishing lines) and interest points (special planar targets) extraction is accomplished.
With relation to vanishing lines extraction, a hierarchical method divided into two levels is applied. In the first level, linear segments are detected applying Canny filter

[Can68] and a linear regression based on least square which combines RANSAC estimator [FB81]; in the second level, segments are clustered through an iterative process based on their colinearity, taking an orthogonal distance as input parameter or threshold. Nevertheless, the presence of mini-segments could carry some problems in the clustering process, i.e. leaving unclassified vanishing lines. In this sense, a weight factor for the line coverage has been considered, which depends on the number of collinear segments as well as their length.

Regarding to the extraction of special planar targets, a seed of the planar target will be required in order to perform a cross-correlation template matching method. The probable target candidates are searched all over the high resolution image using a cross-correlation template matching method. A sub-sampled version of the high resolution image is used to decrease the computational expense. The window size is selected as 10x10 pixels. Only those pixels that have cross-correlation values greater than a predefined threshold value are defined as the target candidates. The algorithm starts searching the most probable target candidates all over the image using cross-correlation values. The seed used in the cross-correlation procedure is generated artificially according to the real target shape (Figure 3). Obviously, the presence of outliers will carry that more targets than the real number will be detected. In this sense, the own radiometric and geometric characteristics of the targets such as green background and circular shape allow filtering some of these anomalies. Finally, with the filtered candidates circular shapes will be extracted through the Generalized Hough Transform (GHT) [Bal81]. This method is a generalization of the traditional Hough transform and allows detecting basic shapes independently of the rotation and scale of the image, event pretty common when we work with oblique images.

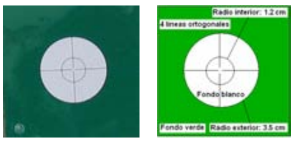


Figure 3: *Special planar targets.*

3.2 Vanishing points computation

The motivation and usefulness of precise and reliable determination of vanishing points, among other structural elements belonging to oblique images, has been demonstrated based on their correspondence with the three main orthogonal directions. Particularly, in architectural environments vanishing points provide three independent geometrical constraints which can be exploited in several ways: from the camera self-calibration and a dimensional analysis of the object to its partial 3D reconstruction.

Our vanishing points method takes a scientific approach which combines several proven methods supported by robust and statistical techniques. In this sense, the key differences of this method in relation with others approaches are reflected in the following steps:

A *Clustering step*, which cluster the mini-segments in vanishing lines.

An *Estimation step*, which applies a modification of the Gaussian sphere method [Bar83], in order to obtain an estimation of vanishing points and reject possible erroneous vanishing lines.

A *Computation step*, which applies a re-weighted least square adjustment support by M-estimators [Dom00].

More details about this new vanishing points method are described in [Agu05].

3.3 Calibration estimation: vanishing points

Our approach is similar to another approach [CT90] who exploiting vanishing points geometry recovers the projection matrices directly. However, in our case the method developed, uses simple properties of vanishing points adding some geometrical constraints derived from image processing step.

The camera model can be recovered following two steps, in which internal and external parameters are estimated separately.

In the first step, the intrinsic parameters, that is, the focal length, the location of the intersection between the optical axis and the image plane and the radial lens distortion, are recovered automatically based on vanishing points geometry and image processing. In the second step, the extrinsic parameters, that is, the rotation matrix and the translation vector which describe the rigid motion of the coordinate system fixed in the camera are estimated in a double process. Firstly, the rotation matrix, that is, camera orientation is obtained directly based on the correspondence between the vanishing points and the three main object directions. This relationship allows to extract the cosine vectors of optical axis, obtaining directly the three angles (axis, tilt, swing). Then, the translation vector, that is, the relative camera pose is estimated based on some priori object information, i.e. a distance, together with a geometric constraint defined by the user. Thus, the reference frame for the camera pose estimation is defined with relation to the object geometry arbitrarily.

The robustness of the method depends on the reliability and accuracy of vanishing points computation, so the incorporation of robust M-estimators in the step before is crucial.

3.4 Dimensional analysis

With the estimation of camera model and with the geometrical constraints defined by the user, an automatic dimensional spatial analysis based on distances and angles is performed between all possible targets combinations. Thus, for each target extracted in image processing step we compute the distances and angles with the remainder targets.

This approach is supported by constrained colinearity condition (3.1) (3.2) and trigonometric functions (3.3), which allow to obtain spatial distances and angles between whatever detected target:

$$\boxed{\begin{aligned} w(1) &= r_{11}(x_1 - x_{PP}) + r_{21}(y_1 - y_{PP}) - r_{31}(f) \\ w(2) &= r_{12}(x_1 - x_{PP}) + r_{22}(y_1 - y_{PP}) - r_{32}(f) \\ w(3) &= r_{13}(x_1 - x_{PP}) + r_{23}(y_1 - y_{PP}) - r_{33}(f) \\ w(4) &= r_{11}(x_2 - x_{PP}) + r_{21}(y_2 - y_{PP}) - r_{31}(f) \\ w(5) &= r_{12}(x_2 - x_{PP}) + r_{22}(y_2 - y_{PP}) - r_{32}(f) \\ w(6) &= r_{13}(x_2 - x_{PP}) + r_{23}(y_2 - y_{PP}) - r_{33}(f) \end{aligned}} \tag{3.1}$$

where, $w(1)..w(6)$ are auxiliary functions derived from colinearity condition, $r_{11}...r_{33}$ is the rotation matrix coefficients, x,y are image coordinates, x_{pp},y_{pp} are principal point coordinates and f is the focal length.

$$\boxed{\begin{aligned} X_s &= X - \left(\frac{w(2)}{w(1)} \cdot \left(\frac{DT_{XZ}}{\sqrt{\left(\frac{w(5)}{w(4)} - \frac{w(2)}{w(1)}\right)^2 + \left(\frac{w(6)}{w(4)} - \frac{w(3)}{w(1)}\right)^2}}\right)\right) \\ Y_s &= Y - \left(\left(\frac{DT_{XZ}}{\sqrt{\left(\frac{w(5)}{w(4)} - \frac{w(2)}{w(1)}\right)^2 + \left(\frac{w(6)}{w(4)} - \frac{w(3)}{w(1)}\right)^2}}\right)\right) \\ Z_s &= Z - \left(\frac{w(3)}{w(1)} \cdot \left(\frac{DT_{XZ}}{\sqrt{\left(\frac{w(5)}{w(4)} - \frac{w(2)}{w(1)}\right)^2 + \left(\frac{w(6)}{w(4)} - \frac{w(3)}{w(1)}\right)^2}}\right)\right) \end{aligned}} \tag{3.2}$$

where, X_S, Y_S, Z_S are the viewpoint coordinates, X, Y, Z are the ground point coordinates and DT is the spatial distance that we want to compute.

With relation to trigonometric functions, for a triangle in the Euclidean plane with edges a, b, c and opposite angles α, β, γ, the following holds:

$$\boxed{\begin{aligned} & a^2 = b^2 + c^2 - 2bc\cos\alpha; b^2 = a^2 + c^2 - 2ac\cos\beta \\ & c^2 = a^2 + b^2 - 2ab\cos\gamma \end{aligned}} \tag{3.3}$$

The accuracy of the method taking into account the inherent conditions in a single image-based modeling approach is around ±10 cm. Nevertheless, this is not especially crucial since we consider that special targets are enough separate each others. So, in most of the cases, a global approximation is usually enough for a search of correspondences.

3.5 Image registration: matching correspondences

This step presents a technique to perform an automatic matching between 3D and 2D points (special targets) belonging to laser model and high-resolution images respectively.

The solution that we propose is based on the invariants properties of two distances and one angle, which are translational and rotational invariant parameters independently of the sensor viewpoint. Furthermore, three of the angle/distance elements, in which at least one of them must be distance, can exactly define a triangle. Therefore, the presented search scheme is the same as to find the equal 3D triangles in both point sets. This search will serve also for rejecting possible outliers. Those points whose correspondence of invariants or triangles is not found of will not be considered. In the end, a final list with the correspondences of target points will be obtained which will constitute the input data in the calibration tuning process.

The method developed for establishing correspondences between both datasets relies on the approach developed by [Akc03], who in order to materialize the correspondence between two laser scanner datasets develops a search of invariants supported by two distances and one angle. Nevertheless, Acka works directly with two homogeneous datasets, which proceed to the same sensor, and with a previous measurement of the invariants obtained through surveying techniques. In the approach presented here, a correspondence between two heterogeneous datasets (2D image points and 3D laser points) is established.

In order to search homologous points, all possible space angles and distances are calculated in both datasets, one through the single image-based modelling approach proposed before and the other directly through 3D coordinates extracted from laser points cloud. The total computational cost for the distances and angles is given below (3.4):

$$\boxed{C\binom{Ni}{2} + C\binom{Nc}{2} + Ni \cdot C\binom{Ni-1}{2} + Nc \cdot C\binom{Nc-1}{2}} \tag{3.4}$$

where, *Ni* is the number of points in the candidate image target list and *Nc* is the number of points in the laser target list, and *C* stands for the combination operator.

Every space angle and two distance combinations for each point in the image target candidate list is searched in the target laser list with a predefined angle/distance threshold values (i.e. angle <0.5º, distance <15cm). Three of the angle/distance elements, in which at least one of them must be distance, can exactly define a triangle. Therefore, the presented search scheme is the same as to find the equal 3D triangles in the both point sets. If a point does not have a compatible 3D triangle in the invariants list, this point does not have a label, namely this point as a wrong target candidate, and must be discarded from the candidate target image list.

3.6 Calibration tuning: image re-projection

A Tsai camera calibration technique [Tsa89] is used to obtain a calibration tuning, especially the image registration with relation to laser scanner. Its implementation needs correspondences between both datasets: 3D laser points and 2D image points. Tsai's technique uses a two-stage approach to compute: first, the position and orientation and, second, the internal parameters of the camera. Thus, Tsai's approach offers the possibility to calibrate internal and external parameters separately. This option is particularly useful in our case, since the single image-based modeling method developed allows us to known these parameters with a similar strategy.

Depending on internal parameters accuracy, we carry out one or two stage approach of Tsai's camera calibration. So, if good accuracy has been achieved through single image-based modeling method, a minimal number of 5 points will be used to compute the camera pose. Furthermore, the three known rotations angles perform as initial approximations in the algorithm.

Due to the different nature of the sensors, as well as the own characteristics of the single image-based modeling approach, a single run of the algorithm can lead to a camera registration that is not fully satisfactory. To improve the accuracy and reliability of the calibration process, an iterative procedure has been introduced. In this sense, each 3D point detected as special target in the points cloud will be re-projected over the image based on colinearity condition principles and the computed camera parameters. Small discrepancies remain between the projected 3D points and the original extracted image points. The 3D coordinates of the laser scanner and the re-projected corresponding image points constitute the input to compute a new calibration. This iterative process follow until the Euclidean distance between the re-projected points and the original image targets points will be minimized (threshold distance). The general idea is that at each iteration the distance between the two datasets is reduced, allowing a better computation of camera parameters.

To ensure the convergence of the algorithm and the improvement of the initial camera model estimation, the calibration error of each correspondence is computed and recorded. In each new iteration, only matching pairs for which the calibration error decreases are updated, and the other are kept unchanged. In this stage, no robust estimation is used since the step before ensures that no outliers are present within the correspondences.

After the calibration tuning procedure based on this technique, a full model for the camera with relation to laser scanner is available and ready to map textures.

4. Experimental results

We have validated our approach on several different datasets, but we only present the experimental results tested over an emblematic romanic church situated in Avila (Spain), San Pedro's church (Figure 4).

Figure 4: *Original image with special targets (3008x2000 pixels)*

After applying the image processing step, we obtain the different features extracted with sub-pixel accuracy (Figure 5).

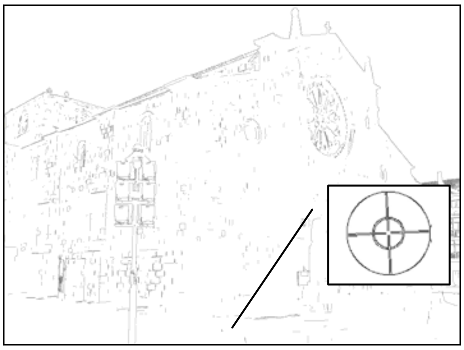


Figure 5: *Image features extraction.*

With relation to the features statistics (Table 2):

Vanishing Lines	231 segments clustered in X direction 274 segments clustered in Y direction 35 segments clustered in Z direction 91 segments clustered as outliers
Special Targets	20 targets were detected 2 targets were not detected 7 targets were detected as outliers
Accuracy (σ)	0.5 pixels

Table 2: *Statistics in features extraction.*

Next, a robust method for vanishing points computation which combines Danish M-estimator together with Gaussian Sphere was applied iteratively (Table 3).

Vanishing points computation (4th iteration)

Gauss Sphere +Danish est. (Unit: pixels)	**VPX**	**VPY**	**VPZ**
x	3761.981	-1483.7	1054.8
y	1432.395	1255.76	-2378.6
σ_{xx}	0.13	0.35	0.38
σ_{yy}	0.18	0.40	0.57

Table 3: *Vanishing points computation.*

With the structural support provided by vanishing points an estimation of camera calibration parameters was obtained (Table 4):

Camera calibration estimation: vanishing points

Internal Parameters (Unit: millimetres)		**External Parameters** (Unit: degrees, metres)	
PP [x]	11.83	**Axis**: 38.00	**X**: -14.95
PP [y]	7.76	**Tilt**: 95.80	**Y**: -12.03
Focal	18.10	**Swing**: 181.44	**Z**: 1.7
K_1	0.003245		
K_2	-0.00001		

Internal Parameters (Unit: millimetres)		**External Parameters** (Unit: radian, metres)	
σ_{PP} [x]	0.032	σ_{Axis}: 0.00175	σ_X: -0.034
σ_{PP} [y]	0.036	σ_{Tilt}: 0.00213	σ_Y: 0.039
σ_F	0.044	σ_{Swing}: 0.00127	σ_Z: -0.048
σ_{K1}	0.000134		
σ_{K2}	0.000001		

Table 4: *Camera calibration estimation.*

Dimensional analysis: single image-based modeling

Taking into account the threshold fixed to distances and angles: 15cm and 0.5º respectively, every space angle and two distances combinations for each point in the image target candidate list was searched in the target laser list, obtaining the following:

- 7 correspondences were located between both datasets and added to the target image list.

- 6 especial targets were detected as outliers being discarded from the candidate target image list.

In the laser scanning context, all special targets were correctly extracted by laser scanner software, so eleven 3D points were added to the laser points list (Figure 6).

Figure 6: *Laser model and the special targets extracted.*

Finally, once both datasets were matched each other in image and laser list, a camera calibration tuning with seven correspondences was performed in order to provide an image to laser model registration (Table 5).

Camera calibration tuning: image to laser model registration

Image to laser model registration (5th iteration) (Units: degrees, radian and metres)	
Axis: 38.6257; σ_A: 0.00065	**X**: 8.138; σ_X: -0.009
Tilt: 95.667; σ_T: 0.00023	**Y**: -3.307; σ_Y: 0.012
Swing: 181.9487; σ_S: 0.00077	**Z**: 1.094; σ_Z: -0.021

Table 5: *Calibration tuning: image registration.*

A re-projection strategy (section 3.6) based on five iterations was necessary to minimize the Euclidean distance between matched points, obtaining good results in mapping textures (Figure 7).

Figure 7: *Multi-sensor registration: mapping textures.*

5. Conclusions and future perspectives

We have developed a method for registering high-resolution images to laser models. Our technique uses a single image-based modeling approach which provides relevant data: from camera calibration and geometrical constraints to a metric dimensional analysis. Particularly, in the automatic co-registration of terrestrial laser scanner data and single digital images, our approach performs a dimensional analysis from a single image based on a search of spatial invariants: two distances and one angle. This approach works very well for outdoors scenes in which the geometry of the building is easy to modeling. Nevertheless, some ill aspects have been assessed: In the search of correspondences step, maybe applying an adaptative threshold supported by a RANSAC estimator could be a good idea to reject fewer points. Obviously, a large sensor's baseline does not contribute in a good way to map textures, obtaining some anomalies in upper parts of the building.

With relation to future perspectives, the research could be extend to exploit the single image-based modeling towards applications related with the improvement and refinement of the laser model, adding metric and semantic information in missing areas (non reflective material, occlusions, shadows, etc). Furthermore, in the context of texture mapping, develop algorithms that allow to handle the resulting problem of occlusions, illumination properties and transition between junctions, would let to achieve a realistic and integral representation of the object.

References

[Agu05] AGUILERA D. G., 2005. Doctoral Thesis. University of Salamanca.

[Akc03] AKCA DEVRIM, 2003: Full automatic registration of laser scanner point clouds. ETH Zurich.

[ATS*03] P. K. ALLEN, A. TROCCOLI, B. SMITH*. New methods for digital modeling of historic sites. IEEE

[Bal81] BALLARD D. Generalized Hough transform to detect arbitrary patterns, 1981. IEEE, 13(2):111-122.

[Can86] CANNY J.F. 1986. A computational approach to edge detection. IEEE journal.

[CT90] CAPRILE B. and TORRE V. 1990. Using vanishing points for camera calibration. IJCV 4 (29).

[Dom00] DOMINGO PRECIADO ANA, 2000. Tesis Doctoral. Universidad de Cantabria.

[Els98] M. D. ELSTROM. Master's thesis, University of Tennessee, Knoxville, 1998.

[FIM*05] FINAT CODÉS J*: The roman theatre of Clvnia: hybrid strategies. ISPRS Workshop Italy.

[FB81] FISCHLER, M. A., and R. C. BOLLES, 1981. Random sample consensus. CM Communications.

[INN*03] IKEUCHI, NAKAZAWA,*. Creating virtual buddha statues through observation. IEEE Workshop.

[LHS01] H. P. LENSCH, W. HEIDRICH*. A silhouette-based algorithm for texture registration and stitching, Graphical Models, vol. 63, n4.

[LPC*00] M. LEVOY, K. PULLI, B. CURLESS,*, The digital Michelangelo project: 3D scanning of large statues. Siggraph, CG Proceedings, 2000.

[MNP*99] MCALLISTER, L. NYLAND, V. POPESCU,*. Eurographics Rendering Workshop, 1999.

[RCM*99] C. ROCCHINI, P. CIGNOMI, C. MONTANI,*, "Multiple textures stitching and blending on 3D objects,". Eurographics, New York, 1999.

[SA01] I. STAMOS and P. K. ALLEN. Automatic registration of 2-D with 3-D imagery in urban environments. Proceedings ICCV-01, 2001.

[Tsa89] TSAI Y., 1989. Synopsis of recent progress on camera calibration for 3d machine vision. Robotics Rew.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Co-registration of Photogrammetric and Laser Scanner Data for Generation of 3D Visual Models

K. Al-Manasir and C. S. Fraser

Department of Geomatics, University of Melbourne, Australia

Abstract

Nowadays, most terrestrial laser scanner (TLS) systems provide the facility of mounting a digital camera on the laser scanner. This not only facilitates a means to generate better quality photorealistically textured 3D models from TLS point clouds; it also offers the opportunity of using photogrammetric orientation techniques to complement existing methods of TLS point cloud registration. This paper describes an approach whereby a registration procedure based upon photogrammetric means is employed as the first step in integrating TLS data and imagery for the generation of textured 3D models. The approach, called image-based registration (IBR), entails an estimation of transformation parameters between the individual scan data and between digital imagery using photogrammetric bundle adjustment. Once both TLS and photogrammetric data are registered in the same coordinate system, the process of forming a segmented structured surface model and its associated triangular mesh are carried out. Photogrammetrically derived constraints are used to convert the unstructured, registered laser scanner model to a structured model. Finally, texture mapping takes place via the rectification of image patches from the integrated images used in the IBR process onto individual surface elements. Test results obtained with the proposed approach are presented to highlight its practicability and accuracy.

1. Introduction

Digital imagery is now being routinely used in conjunction with terrestrial laser scanner (TLS) data to generate photorealistic 3D object and scene models, since the use of recorded laser scanner intensities alone generally produces a texturing of insufficient visual quality, especially when there is other than a very high resolution scan. TLS manufacturers generally offer the option of having an SLR-type digital camera mounted on the scanner. This gives rise to the necessity of establishing the interior and exterior orientation of the camera such that the position and attitude of each image can be established with respect to the XYZ coordinate system of the TLS.

TLS scans can be registered in one coordinate system using the ICP algorithm [PJND92], which requires pre-alignment of the individual TLS point clouds, although automated registration of TLS scans is also possible without the requirement for initial values for the transformation parameters [Dol05]; [NC05]. A common alternative registration approach is to position automatically recognisable artificial targets such that they form common or tie points between adjacent point clouds thus providing registration through 3D coordinate transformation.

With respect to imagery from the camera mounted on the laser scanner, the relative orientation of adjacent pairs of images, and indeed of a network of overlapping images, can be accomplished by measuring conjugate image points, either manually or automatically in cases where either special targets are used or where the geometry is such that image matching is feasible. Registration of the photogrammetrically derived 3D coordinates and the TLS point cloud is also achieved if the exterior orientation of the camera stations is known with respect to the TLS. Once the 3D model and the digital imagery are registered in the same coordinate system, the next step is to convert the point cloud into a structural surface, usually in the form of a triangular mesh. Finally the texture mapping can take place and colour patches from the imagery can be assigned to each individual triangular facet in the 3D model.

The scenario considered in this paper for the generation of a photo-realistically textured 3D model via the use of a TLS with attached digital camera is the following:

- In the first stage, scanning occurs at each station and digital imagery is recorded from the camera mounted on top of the laser scanner, so that the scene being scanned is also photographed. Since the camera likely has a limited field of view compared to the TLS, it is possible that some areas covered by the scan will not be imaged.
- After the scanning is completed, the camera is removed from the TLS and supplementary images are recorded to make sure that the entire object is covered.
- The TLS and photogrammetry point clouds are registered in the same coordinate system using the Image Based Registration (IBR) technique, first proposed in [AF06a]. The IBR registration provides a photogrammetric approach to point cloud registration. Images from the TLS-mounted digital camera are first used to relatively orient the network of images, after which the exterior orientation between TLS point clouds is determined based on the known relationship between the position and orientation of the camera and TLS.
- The photogrammetric network established for the IBR is supplemented by any additional images recorded using the camera dismounted from the TLS.
- Lines and planes are triangulated using the photogrammetric network. These photogrammetrically extracted features are then used as constraints to improve the segmentation process for the point cloud. They are also applied in the mesh generation process to preserve the geometric characteristics of the model.
- Finally the 3D model is segmented and converted into a triangular mesh and the texture mapping takes place using image patches from appropriate images forming the full photogram-metric network.

Each of these stages will now be discussed.

2. Image-based registration

The image-based registration method for TLS scan data is fully described in [AF06a] and a short summary of the process only will be presented here. First, a camera calibration is required. Second, the camera position with respect to the laser scanner coordinate system must be recovered via spatial resection. For a rigidly mounted camera, this process need only be carried out once. Finally, once the TLS and image data from two or more scenes are recorded, the registration process can be carried out. Only the imagery must overlap; there is no requirement for the TLS point clouds to overlap.

The initial step of the IBR is an analytical relative orientation between two adjacent images using five or more suitably located conjugate points, which can be automatically detected and measured in cases where image-identifiable coded targets are employed. Whereas the TLS will be continually collecting laser range data throughout its lateral sweep, only one image needs to be recorded, with the requirement that there is suitable overlap with the image from the second station. From the five relative orientation parameters determined (three rotations and two translations) and from knowledge of the transformation parameters between the camera coordinate system and the TLS, the exterior orientation and hence the registration of the laser point clouds is established, without the need for point correspondence searching or an ICP registration procedure. In the situation where there are more than two overlapping images and TLS point clouds, a bundle adjustment can be applied for the initial multi-image orientation.

Once the relative orientation between two images i and j is determined, the point cloud registration can be established using the following equation:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_i = \mathbf{R}_{\mathrm{A_i}}^{-1} \cdot \left(\mathbf{R}_C^{-1} \cdot \left[\mathbf{R}_{j,i}^{-1} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}_j + \begin{pmatrix} bx \\ by \\ bz \end{pmatrix} \right] + \begin{pmatrix} X^c \\ Y^c \\ Z^c \end{pmatrix} \right) \quad (1)$$

where

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_j = \mathbf{R}_{\mathrm{C}} \cdot \left[\mathbf{R}_{\mathrm{A_j}} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_j - \begin{pmatrix} X^c \\ Y^c \\ Z^c \end{pmatrix} \right]$$

Here, i is the reference point cloud and j is the data set whose coordinates are to be transformed; the 3x3 rotation matrix $\mathbf{R}_C$ and the translation vector (Xc, Yc, Zc) express the camera position and orientation in the TLS coordinate system at a specific alignment of the scanner; the 3x3 rotation matrix $R_{j,i}$ and the vector (bx, by, bz) are formed by the exterior orientation of camera station j within the coordinate system of camera station i; and R_A is a 3x3 rotation matrix defining the TLS rotation around its Z-axis described by the scanner rotation angle A at the time of exposure. The accuracy of the registration can be enhanced through the addition of extra images (without TLS data) in the bundle adjustment. Experimental evaluations of the IBR approach have been reported in [AF06a] & [AF06b].

3. Integration of additional image data

The additional imagery recorded with the camera dismounted from the TLS can be integrated into the IBR-registered network via either photogrammetric means alone, ie added to the bundle adjustment or in the less likely case where there is insufficient image overlap, via spatial resection from the registered point cloud. At this stage it is also appropriate to carry out any point cloud preprocessing such as a point density reduction and outlier detection, especially in the overlap areas of the TLS point clouds where there will be redundant data [Rem03], Gaps in the point cloud can subsequently be filled manually, semi-automatically or automatically. Point cloud decimation, outlier detection and gap filling processes all use surface gradient information for the preservation of the geometry (shape) of the model.

4. Photogrammetrically-derived constraints

The presence of very dense laser point clouds can complicate and adversely affect the accuracy of feature point identification and segmentation. Also, edge identification and definition is influenced by scan point density. As an aid in rectifying such problems, photogrammetrically derived constraints can be applied within the 3D Delaunay triangulation to yield a more accurately structured model. Photogrammetrically extracted features such as lines and planes are used as constraints both to improve the segmentation process for the laser point cloud and to preserve the geometric characteristics of the model.

5. Segmentation and reconstruction

After the scan data is transformed into a common reference system, the registered 3D model should be further processed to convert the unstructured point cloud into structured form, usually via predefined elemental primitives that represent the object. These include best-fitting edges, planes, spheres, planar facets and cylinders. This process, in which points with the same homogeneous properties are grouped into regions, is referred to as segmentation and surface fitting. The segmentation process can be divided into two categories, namely surface-based and edge-based segmentation. In the first, the segmentation process is based on point clustering for surface shape representation, whereas in the second, the process utilises discontinuities within the data.

Use of the surface normal to group points in clusters has been reported in [RAK87] and various methods for range data segmentation are summarised in [PR88], where an iterative region growing method for surface segmentation is also reported. In this method, the mean and curvature of a point cluster are used to group the data, curvature being invariant to rotation and translation of the coordinates. The solution is iterated and a best-fit surface is estimated until a threshold value is achieved.

6. Texture mapping

Once the object model has been completely created in term of a triangular mesh or via a parametric approach, the final step of the generation of a visually realistic digital model is the mapping of texture onto the individual surface elements. This can be achieved by mapping 2D image elements, with appropriate rectification, onto planar surface elements of the 3D object model. The following principal stages are involved in the texture mapping for each surface element (polygon or triangle) [SFCG98]:

1. Selection of the appropriate image from the set of images in which the surface element appears,
2. Determination of 2D image coordinates for the points forming the surface element from the corresponding 3D object space coordinates and the exterior orientation of the image,
3. Specification of 3D and texture coordinates in a given modelling language such as VRML, and
4. Viewing the scene using a standard viewer.

In a multi-image network, there will likely be sufficient overlap between images such that a triangular facet in object space will appear in a number of images. Thus, several textures for the surface element will be possible. As one solution to this problem, the texture is selected from the image where the triangle appears largest. However, this may result in discontinuities in the adjacent triangles. In order to reduce the discontinuities in texturing which accompany the use of multiple images, a weighted averaging approach can be adopted [SFLM*03].

7. Experimental modelling of Cooks' cottage

7.1 TLS and photogrammetric recording

As a test of the proposed approach, a TLS survey of Cooks' Cottage, a heritage site and popular tourist attraction in Melbourne, was carried out with a Riegl LMS-Z210 / Nikon D100 scanner/camera combination, with the geometry shown in Figure 1. Four images where recorded with TLS scans (TLS i) and four additional images were recorded with the camera removed from the scanner. Care was taken to ensure that there would be sufficient overlap between images to support robust relative orientation and subsequent bundle adjustment. However, little attention was paid to the extent of TLS point cloud overlap.

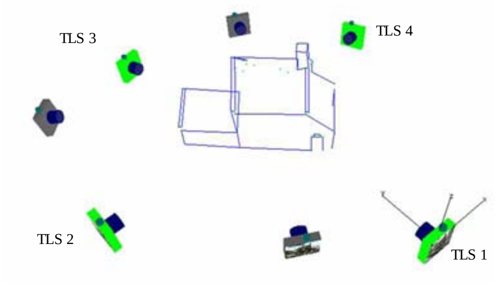


Figure 1. *Scanner/camera station geometry for the survey of Cooks' cottage. (Images from the TLS-mounted camera are labelled TLS i)*

A relative orientation was first performed between the images from stations TLS 1 and 2. This was followed by an initial resection of the images at stations TLS 3 and 4, and of those from the supplementary camera stations. A bundle adjustment, with an average of 30 points per image, was then performed using all seven images This produced an RMS value of image coordinate residuals of 0.4 pixels and an estimated point positioning accuracy of 2 mm. Registration of the TLS point cloud data was then carried out via the IBR method using Eq. 1, with the resulting registered 3D model being shown in Figure 2. Following the IBR process, a registration using the ICP algorithm was also performed to produce a second 3D data set.

The accuracy of the registered 3D coordinates obtained with the IBR was verified using the coordinates of 120 well distributed photogram-metrically measured checkpoints of 2mm accuracy. These were manually identified in both the IBR & ICP generated 3D point clouds. The resulting RMSE values for the ICP and the IBR models, as assessed against the checkpoint coordinates, were 4mm and under 3mm, respectively. Given both the accuracy of the Riegl LMS-Z210 and the limited ability to precisely identify the checkpoints in the laser data, the results are consistent with expectations, though it is noteworthy that the IBR produces higher accuracy than the ICP approach. It must be remembered, however, that the test survey was specifically designed to produce a sub-optimal ICP solution, since an aim of the exercise was to show the merits of the IBR in cases where there is low overlap between point clouds from adjacent TLS stations.

Figure 2. *Laser scanned 3D model of Cooks' cottage.*

7.2 Model reconstruction and texture mapping

The triangular mesh-model shown in Figure 3 was created using the previously described constrained 3D Delaunay triangulation. Photogrammetrically derived features, mainly lines and planes, were used to convert the unstructured, registered TLS point cloud into a structured model.

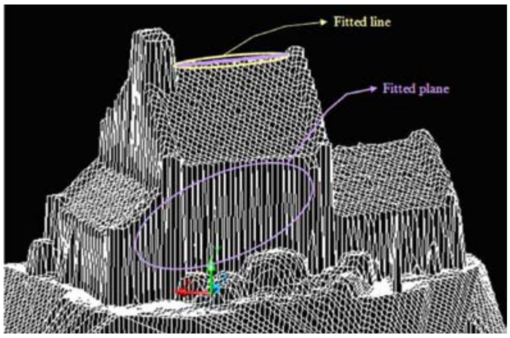


Figure 3. *Triangulated mesh-model of Cooks' cottage.*

Following the reconstruction process, the texture mapping was carried out. Initially, aggregated surface areas such as planar walls and roof sections were assigned a texture via appropriate rectification from the most optimal image. Following this, irregular surfaces were textured on an individual mesh triangle basis, using the method discussed earlier. Views of the final texture-mapped model are shown in Figure 4.

The accuracy of the registered 3D textured-mapped model was quantified using the coordinates of 80 photogrammetrically measured checkpoints of 2mm accuracy. These were manually identified. The resulting RMSE value for the texture-mapped, photogrammetrically-constrained TLS model was 2.5mm.

(a)

(b)

Figure 4. *Textured 3D model of Cooks' cottage.*

8. Conclusions

A process for the creation of 3D texture-mapped models via the integration of terrestrial laser scanner and photogrammetric data has been described. The process commences with the technique of image-based point cloud registration. The step that follows then involves the use of photogrammetrically derived features such as lines and planes to form constraints in the generation of a structured model from the laser point cloud through 3D Delaunay triangulation. Finally, texture mapping is carried out using image patches from appropriate images forming the full photogrammetric network.

Promising results have been achieved with the approach, which can yield higher modelling accuracy in cases where the photogrammetric orientation and triangulation is inherently more precise than the particular TLS system being employed. There is also the advantage of likely better interpretability of object feature constraints in the imagery as compared to the TLS data. Further advantages of the method are that the use of the IBR method does not require any overlap between adjacent laser point clouds to perform the registration process, and that supplementary images can be used to enhance the photogrammetric solution.

References

[AF06a] AL-MANASIR, K., and FRASER, C. S., 2006a. Registration of terrestrial laser scanner data using imagery, *Photogrammetric Record* (in press).

[AF06b] AL-MANASIR, K., and FRASER, C. S., 2006b. Registration of TLS data using imagery. *Proceeding of the ASPRS Annual Conference*. Reno, Nevada, May 1-5.

[PR88] BESL, P., and JAIN, R., 1988. Segmentation Through Variable-Order Surface Fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(2), 167-192.

[PJND92] BESL, P. J., and MCKAY, N. D., 1992. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239-256.

[Dol05] DOLD, C., 2005. Extended Gaussian images for the registration of terrestrial scan data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36 (3/W19), 180-186.

[SFCG98] EL-HAKIM, S. F., BRENNER, C., and ROTH, G., 1998. An Approach to Creating Virtual Environments Using Range and Texture. *Proceedings of the ISPRS Symposium: Real-Time Imaging Dynamic Analysis*, Hakodate, Japan, 331-338.

[SFLM*03] EL-HAKIM, S. F., GONZO, L., PICARD, M., GIRARDI, S., SIMONI, A., PAQUET, E., and VIKTOR, H. B., 2003. Visualization of Highly Textured Surfaces. *4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST2003)*, Brighton, United Kingdom.

[RAK87] HOFFMAN, R., and JAIN, A. K., 1987. "Segmentation and Classification of Range Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5), 608-620.

[Rem03] REMONDINO, F., 2003. From Point Cloud To Surface: The Modeling And Visualization Problem. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIV, part 5/W10 (CD-Rom). *ISPRS International Workshop on Visualization and Animation of Reality-based 3D Models*, Tarasp-Vulpera, Switzerland.

[NC05] RIPPERDA, N., and BRENNER, C., 2005. Marker-Free registration of terrestrial laser scans using the normal distribution transform. *ISPRS Workshop on Virtual Reconstruction and Visualization of Complex Architectures*, Mestre-Venice, Italy.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Towards a Photogrammetry and Virtual Reality Based Heritage Information System: A Case Study of Shawbak Castle in Jordan

Pierre Drap[1] , A. Durand[1] , M. Nedir[1] , J. Seinturier[1] , O. Papini[2] , F. Boucault[3] , P. Chapman[3] , W. Viant[3] , G. Vannini[4] , M. Nuccioti[4]

[1] MAP, umr CNRS 694, Ecole d'Architecture de Marseille, France
[2] LSIS umr CNRS 6168. Université de Toulon et du Var, Toulon, France
[3] SIMVIS, Simulation and Visualization Research Group, University of Hull, Hull, UK
[4] Dipartimento di Studi storici e Geografici dell'Universitá degli Studi, Florence, Italy

Abstract

The paper presents an interdisciplinary project which is the first step towards a 3D Geographical Information System (GIS) dedicated to Cultural Heritage with a specific focus application on the Castle of Shawbak, also known as the "Crac de Montréal" in Jordan.

Current 3D GIS already provide support for urban models on a city scale. Our project however focuses on a building scale encompassing its atomic elements such as ashlars blocks, cement, stratigraphic unit and architectonic elements. At this scale we need a full 3D interface in order to manage accurate measurements and a mainly heterogeneous archaeological documentation.

The project is conducted by four laboratories: the MAP-GAMSAU located in the school of Architecture of Marseilles, France in charge of the photogrammetric survey phase; The LSIS laboratory, France, will be in charge of the knowledge based approach; SimVis from The Department of Computer Science, University of Hull, UK, for the virtual reality aspect and of course the "Dipartimento di Studi storici e Geografici" from the University of Florence, Italy, in charge of the archaeological part.

To manage these archaeological data the project is divided into three phases: The survey phase: using a knowledge based photogrammetric tool, Arpenteur (http://www.arpenteur.net), the photogrammetric campaign ensures a survey founded on archaeological knowledge and directly linked with a database built by archaeologists. The objective here is to link an already existing archaeological database with a photogrammetric tool in order to simplify the photogrammetric process. Our goal here is to offer to the archaeology community a new tool for surveying where technical photogrammetric aspects are more or less hidden from the surveyor. The second phase is the use of the knowledge base to ensure data consistency through a complex and multi-user survey phase. Based on data fusion coming from different sources, this phase will ensure a reversible way to merge several partial surveys exploiting the complementarities between sources, solving different existing conflicts and reducing the possible redundancies. This fusion process deals with archaeological information as well as spatial information. Finally we need a high resolution interface between the final geometry and the archaeological database. Virtual reality using interactive immersive devices and specially designed software tools is an efficient method for revisiting the site and for analysing, updating and revising knowledge.

This project described in this paper is work in progress. After three photogrammetric campaigns in Jordan the first results are available on the project web site: http://www.shawbak.net

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Virtual Reality

1. The archaeological context

The archaeological study is led by the Dipartimento di Studi storici e Geografici of the University of Florence, Italy. The work in Shawbak is part of a wider research aimed at analysing the structural aspects of feudal society all over the Mediterranean basin through a sampling strategy based on 'historical regions' to define spatial contexts. [VNTD02].
One such region is actually the Trans-Jordan of Crusader-Ayyubid age, organised according to western European standards between year 1100 and 1187 (when the Crusader settlement was abruptly dismantled after the defeat suffered by the army of the Latin Kingdom of Jesrusalem).
The settling strategies adopted in the area by king Baldwin I and his followers resulted in the building of large rural fortified settlements (similar to the ones contemporarily created by the feudal aristocracies in southern France or in Italy) located on a line connecting present-day Amman to the red sea.
Such a display of economic and military means was indeed justified with the attempt to control the most important road system of the Arab world (connecting Damascus to Cairo and to the desert 'highways' leading to the Arabian peninsula), and had the ultimate effect to bring back to life the historic frontier of the Roman empire: the so-called limes arabicus. The area spanning from the ancient city of Petra

Figure 1: *A selection of the archaeological sample used to achieve an actual knowledge of the material aspects of feudal society's lifestyles across the Mediterranean basin. All projects are part of the Strategic Research Programme 'The Mediterranean feudal society': archaeological profile is supported and directed by the University of Florence.*

and the site of Shawbak can be considered a real keystone for the Crusader's project, as can be easily demonstrated by the early interests and the specific instructions given by the king himself to organise a settling system right there. The area encloses a number of fortified villages, known in written sources as castra, literally castles. All of them, except one, are concentrated inside (or in the near vicinity of) the urban area of ancient Petra. Two of them (al-Wu'Ayra and al-Habis) have been widely investigated in previous years

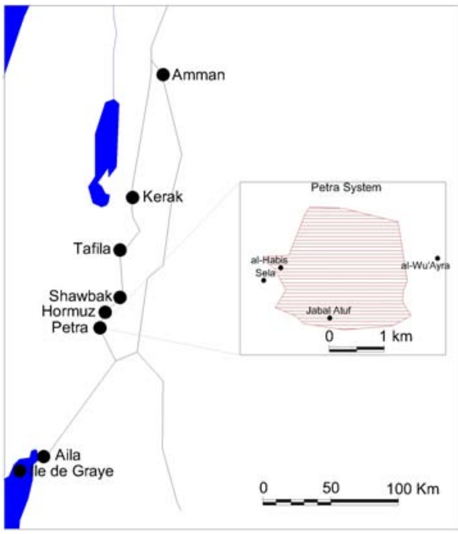


Figure 2: *A schematic view of the European settling system in Transjordan at the beginning of the 12th century.*

by means of traditional and 'light' archaeological means (see [DDS*05] for further details), while a third is currently under study: the castle of Shawbak.
Located approximately 25km north of Petra, the archaeological-monumental site of Mons Regalis/Shawbak can be considered one of the best preserved rural medieval settlements in the entire Middle East. Its key characteristics include a relevant time-spanning readable stratigraphy (from Roman to Othoman periods), an astonishingly well preserved nucleus of still standing medieval historical buildings and (connected with the above) a primary role played over the centuries by the castle (from Crusader age) in the political and military control of the whole Transjordan.
Archaeological readings at the site have always encountered a number of problems relating to data management. In particular, a suitable solution was required that allowed the acquisition, editing and querying in real-time of a large amount of data belonging to different documentary series (i.e. archaeological textual records, archaeological survey, architectural plans/elevations, 3D digital terrain models etc.) so as to maximize the possibilities of historical interpretation. [VN03].
Knowledge based photogrammetry appeared to provide an extremely valuable solution for the envisaged archaeological needs.

2. The 3D survey

2.1. The Arpenteur project

The Arpenteur project (ARchitectural PhotogrammEtry Network Tool for EdUcation and Research) started in 1998 by Pierre Drap, Pierre Grussenmeyer [DG00]. In the past few

years, the project has become both a WEB-based tool and traditional software running in Java on several platforms (Windows, Linux). It has been regularly completed and updated according to the evolution of the Java Development Kit proposed by SUN^{TM}.
Arpenteur is a photogrammetric project devoted to architectural survey that offers a simple and efficient tool for archaeologists and architects and that does not require a deep knowledge or expertise in photogrammetry. Once the first orientation step is performed by a photogrammetrist at least for photogrammetric model control and validation, the measuring step, made with the Arpenteur software, can be performed with the help of experienced researchers of other domains of knowledge, archaeologist or architect. Examples are available at http://www.arpenteur.net.
The project's main objective is founded on the idea of a process guided by knowledge related to one's personal field of study. The results can be shown as documents (XML), visual file (SVG, VRML, X3D) or as a body destined to database. For this purpose, the system makes a set of tools available for the experts and allows them to formulate hypotheses and the corresponding measurements related to their field of investigation.

2.2. A two step Photogrammetry campaign

The photogrammetric campaign was held in November 2004 and was aimed at surveying the fortified gates CF3 and CF5. Raw data of this campaign include: 743 photographs, 234 oriented photographs, 5325 3D points calculated, 100 control points and 10 working days for an archaeologist to make the orientation with Photomodeler. The archaeological team uses two Leica laser total stations that, in a previous field season, were employed to lay out a local reference system for the entire site.

2.3. Measuring 3D blocks

Once all the photographs are oriented The I-MAGE process (Image processing and Measure Assisted by GEometrical primitive), developed in 2001, has been designed to help during the measuring process in photogrammetric surveys. Users can make a 3D measurement using one single photograph, without altering result precision. This method was already published at the CIPA congress, [DDG01]; it allows the user to concentrate on the archeological aspect of the survey with less attention on the photogrammetric one.
We use this approach also to produce 3D models of building blocks (i.e. ashlars) based on the only observable face. The morphology of each ashlars block is expressed as a polyhedron with two parallel sides, or faces. In most of the cases only one side is visible, sometimes two, rarely three. The survey process can inform about the dimensions of one face, then the entire polyhedron is computed accordingly to the architectural entity's morphology (extrude vector) and the data provided by the archaeologist (depth, shape, ...). Computing an extrusion vector can be easy in the case where

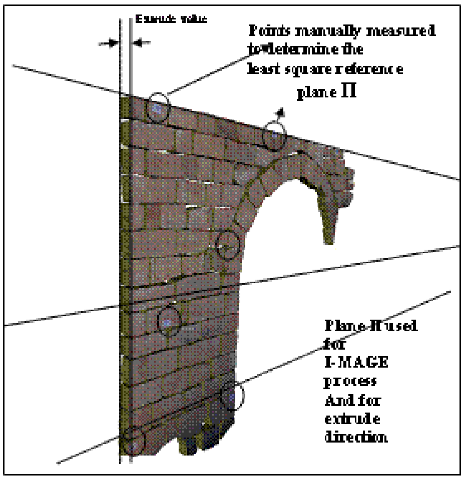


Figure 3: *Extruded ashlars blocks using a plane as an approximation of the exterior face of the wall.*

the architectural entity's morphology is obvious; during a wall survey for example an extrusion vector can be computed by a minor square adjustment of a plane around the survey zone. In our case it's the plane used in the I-MAGE. In this case where the entity's geometrical properties are simple, the extrusion vector is calculated before the survey phase and the block is extruded directly from the measured points. In the case of the survey of an arch the extrusion should be radial and needs the geometrical features of the entity (intrados, radius, axis).
This approach for measuring blocks was already published in a VAST congress [DHVG00] and has been combined with the I-MAGE process in order to obtain an integrated tool.

3. Archaeological data and 3D survey: a single database

This project deals with heterogeneous data. We can group data into two kinds. Archaeological data are made by documentation, conceptual knowledge and classifications. Surveyed data is gathered from photogrammetry, observation and measurements.

3.1. The plotting interface

After the orientation phase done with Photomodeler all the oriented photographs are stored in the database with all the associated computed parameters. The archaeological plotting phase is done with a specific photogrammetric module, using only two images. At this stage the accuracy is sufficient with a measure done with two images and the interface is simpler to manage. The user has to choose two photographs, already measured blocks are displayed and Arpenteur will generate a correspondent photogrammetric model

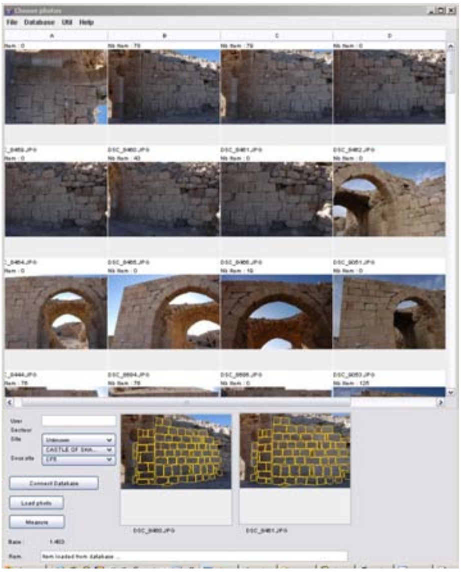


Figure 4: *Choosing oriented photographs from the database with Arpenteur.*

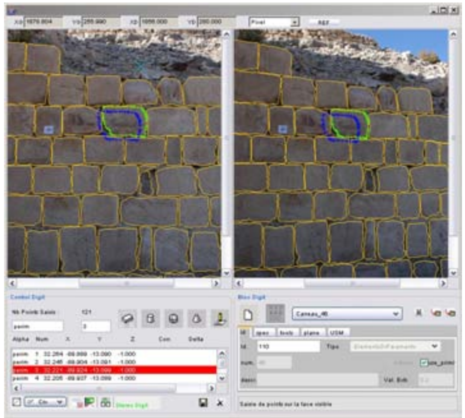


Figure 5: *Arpenteur. Block measuring interface.*

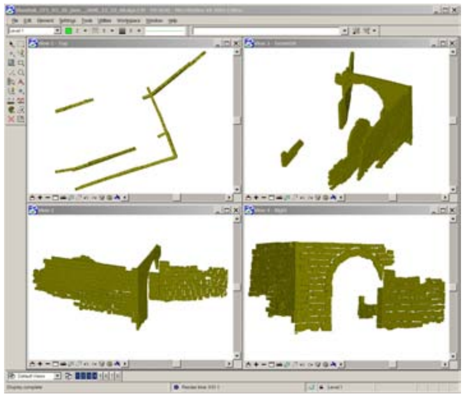


Figure 6: *Shawbak plotting with Arpenteur, exported in CAD Microstation (Bentley).*

on the fly. The application will connect to the database over the Internet to display thumbnail and load photographs and plotted blocks. A direct link to the PHP Database interface is available by picking the displayed blocks. The specific plotting interface, connected to the database was designed according to archaeologist needs and in connection with I-MAGE approach.

The plotting phase result is stored as a set of XML files before they are stored within the database. As several archaeologists can have concurrent access to the database and the survey can be done on several months. In order to ensure a coherent database inconsistency is solved before the final storage in the database. These aspects will be discussed in section 3.3. The Block plotting Arpenteur interface ensures a full integration of the I-MAGE process and allows the archaeologist to perform 3D measurements on two photographs focussing on the block's details.

A link with the database through a PHP based interface allows access to all the archaeological data already stored, neighbouring blocks, stratigraphic units, etc.

3.2. A relational database for archaeological and georeferenced data

Until now the formalism of data storage used in Arpenteur was XML. The constraints and the evolutions which presided over the choice of a relational database storage was initially the growing quantity of the data of the Shawback project, with XML storage becoming heavy in terms of performances and management. Then the choice was to migrate on the SGBD MySql up to version 4.2 which should support foreign keys, transaction and concurrent access. The second main reason for this migration was that the archaeologists of University of Florence developed a database to describe the stratigraphic units, in terms of archaeological knowledge, geometric, and spatiotemporal data. The archaeologists then formulated the need to have a tool which also integrates the data coming from Arpenteur. In the objective to have a unified model facilitating the access and the archaeology data analysis, the choice of a relational database storage was imperative.

The design of this database was conceived to be generic and reusable for all the projects, for all types of objects measured under Arpenteur, and to respect the Arpenteur object model. This was made by storing the Java class in the data description of the measurable item. This item inherits from a generic item which is instantiated by loading common data. The inherited item is then instantiated by loading the items specific data like the item description, the item metrology etc. All the measurement data of 2D and 3D points, photos and models that were used for the orientation are all stored in the database. The restitution of the specific model item is made in the air. The data access object is developed in Java and use a JDBC driver. The JDBC access data object implements usual requests like loading the database for instantiation of the Arpenteur model, insert new item, delete and update an existing item, and requests returning associations between items and photos on which they were measured.

This new data access model was applied on the Shawback project and the genericity was tested by using the same database scheme, the same JDBC access data object, and the same viewing interface for these projects. We need to implement only the specific Arpenteur object interface for each types of measurable item, to have a new application of Arpenteur. This is the main advantage of the Object model.

Our goal now is to extract knowledge based on spatial requests needed by the experts. Some of them are already available like loading all the items located in a sphere or in a bounding box. But close work and collaboration with the experts is required in order to identify their needs and produce an efficient spatial exploration tool. From this perspective, the idea to develop a 3D GIS tool integrating all these data was naturally imperative to provide an efficient research data analysing tool.

Work is beginning to redesign the existing archaeological database in this direction. We will adopt the Harris matrix and the ontology methodology for knowledge acquisition.

The development will be done using Geotools Java interface,(http://geotools.codehaus.org/) with a full integration in Arpenteur, virtual reality tools, and a Java 3D module to define the projection plan. These new research developments will be documented in future papers.

3.3. Using knowledge to ensure consistency

A detailed photogrammetric survey, stone by stone, will generate a huge quantity of measured objects. The plotting phase is extremely time consuming (several months and more than one operator may be required). In addition the data managed is strongly heterogeneous, all these factors increase the probability of inconsistencies in the measured objects set.We propose in this part a formalisation allowing detection of inconsistencies between sets of measured objects as well as an automated treatment of re-establishment of consistency and fusion.

3.3.1. Consistency of a set of objects

Our model of measurable objects relies on the object paradigm associated with a description. It is composed of:

1. The Object Class;
2. Default attributes values;
3. The set of constraints on the attributes denoted by $\mathcal{C}_a$;
4. The set of relations between objects denoted by $\mathcal{R}$;
5. The set of constraints on the relations of $\mathcal{R}$ denoted by $\mathcal{C}_r$.

Aspects 1, 2 and 3 of the model form the intrinsic part. Indeed, these three aspects relate to only the objects themselves. Aspects 4 and 5 form the relational part of the model. Aspects 2, 3 and 5 form the description added to the object formalism. Practically, this description is an XML file assigned to a class. The following example illustrates this model:

Exemple 1 Let $O = \{b_1, b_2, b_3, b_4\}$ be a set of measured objects of the class Block. For the sake of simplicity, we use in this sample only the attributes *name*, $length$, $width$, $volume$, $localisation$ and the set of photographs, denoted by *photos*, on which a block is measured. Constraints on attributes are:

- $0 < b.length < l$
- $0 < b.width < w$
- $0 < b.volume < v$
- $card(b.photos) > 0$

Where l, w and v are values given by the experts. For two blocks b and b', a sample set of relations $\mathcal{R}$ is defined by:

- $sameName(b,\ b')$ iff $b.name = b'.name$
- $sameLoc(b,\ b')$ iff $b.localisation = b'.localisation$
- $commonPhoto(b,\ b')$ iff $b.photos \cap b'.photos \neq \emptyset$

The computation of the relations is not always simple. For example, the computation of the relation $sameLoc(b,\ b')$ have to take into account specific geometry and measurement threshold. Experts given constraint on relation is for example:

- c_l: Two blocks b and b' of O, should not verify $sameName(b,\ b')$ if $sameLocation(b,\ b')$ is not verified.

In this model, a set of objects is consistent if and only if each object satisfies all the constraints on its attributes and each relation of $\mathcal{R}$ satisfy the constraints on the relations. If a set O of object is not consistent, there is a finite number of subset of O which are inconsistent. These sets are distinguished according to the type of constraints which they violate. A set of objects violating the constraints on attributes is called an intrinsic inconsistency set. A set of objects violating constraints on the relation is called a conflict. We now propose a formalism in order to represent our model.

3.3.2. Logic formalisation

In order to formalise the coherence of a set of objects, we use first order logic instantiated on the domain made by the set of measured objects. This choice is carried out by its simplicity of expression as well as its decidability. For us, the domain is the set of measured objects. Usual logical connectors are used $\vee$, $\wedge$, $\neg$, $\leftarrow$ and usual quantifiers are $\exists$, $\forall$. Unary predicates correspond to constraints on attributes and are such that:

Definition 1 Let o be an object of the set O belonging the class A. A constraint denoted by c_a from $\mathcal{C}_a$ enable us to define the unary predicate $c_a^A(o)$ such that:

$c_a^A(o)$ is satisfied if and only if c_a is verified for the object o.

The constraints of the sample 1 are expressed by: $length^{Block}(b) : 0 < b.length < 1{,}50m$, $width^{Block}(b) : 0 < b.width < 1m$, $volume^{Block}(b) : 0 < b.volume < 2m^3$, $photos^{Block}(b) : card(b.photos) > 0$ The set of the unary predicates is denoted $\mathcal{PC}_a$. Binary predicates are attached to relations between objects and are such that:

Definition 2 Let o and o' be two objects of O and let r be a relation of $\mathcal{R}$. The binary predicate $r(o, o')$ is defined by:

$$r(o, o') \text{ is satisfied if and only if } o \text{ et } o' \text{verify } r.$$

Predicates representing relations of the sample 1 are: $sameName(b, b')$, $sameLoc(b, b')$, $commonPhoto(b, b')$, $pointsMix(b, b')$. Finally, constraints on the relations are represented by formulas whose terms are binary and unary predicates. The constraints on relations for the sample 1 are:

$$c_l : \forall b, b' \in O \neg sameName(b, b') \wedge sameLocation(b, b')$$

The set of binaries predicates is denoted by $\mathcal{PC}_r$.

3.3.3. Consistency of a set of objects

This formalisation enables us to define the coherence of a set of objects.

Definition 3 Let O be a set of measured objects and let $\mathcal{PC}_a$ and $\mathcal{PC}_r$ be the set of predicates representing the model constraints. O is consistent if and only if it is model of the set of formulas $\mathcal{PC}_a \cup \mathcal{PC}_r$.

If a set of objects is inconsistent, the inconsistent subsets are needed. Each one of these subsets is associated with the constraints which it violates to form a state. More formally, a state may be described as:

Definition 4 Let O be a set of measured objects and let O' be a subset of O violating the set of constraints $\mathcal{PC} \subset \mathcal{PC}_a \cup \mathcal{PC}_r$. A state , denoted by e is the couple $(O', \mathcal{PC})$.

The state associates the inconsistent subset of measured objects the violated constraints, and thus the cause of the inconsistency. We are currently working on an application which enables us to provide statistics to an expert and make treatments to solve inconsistency. This application will rely on the latest logical solver to find inconsistency.

4. An immersive virtual reality interface for Shawbak Castle

In recent years, VR has greatly facilitated the interaction and interpretation of archaeological data. For example the remarkable work of Vote and her team [VALJ02] who provided archaeologists with an efficient tool to interact with the Great Temple in Petra, Jordan using a CAVE virtual environment.

We are using the Hull Immersive Visualization Environment (HIVE) facilities [PCVH04] in order to provide a 3D GIS interface to the relational database described above. The HIVE includes an auditorium with a large screen, stereo video projectors and a Vicon tracking system.

Figure 7 provides a high level description of the HIVE architecture used in this project. It shows the rear projected wall (thus eliminating shadows), main and tracking computer, infrared cameras for the tracking systems, a user wearing shutter glasses for the stereo display and a special pointing device for interacting with the Shawbak castle database.

Our rendering system has been developed using *OpenSceneGraph*. *OpenSceneGraph* (OSG) is an open source high performance 3D graphics toolkit, used by application developers in fields such as visual simulation, games, virtual reality, scientific visualization and modelling. OSG is written in C++ and OpenGL and runs on all Windows platforms, OSX, GNU/Linux, IRIX, Solaris and FreeBSD operating systems. The geometry acquired with

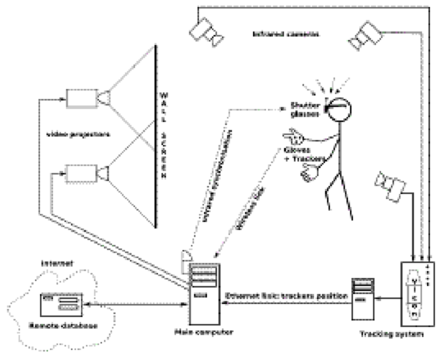


Figure 7: *HIVE architecture.*

the I-MAGE photogrammetry technique is retrieved from the MySQL database at the start of the application. The geometry acquired is cached which makes future reloading instantaneous. The user is able to dynamically choose which blocks they want to display at any time: the software maintains a continuous connection to the database and requests blocks required that have not yet been loaded into memory.

The drawing of the blocks is an extrusion of the polygon retrieved from the database. The extrusion value will have been previously specified and entered into the database by archaeologists. Finally, the blocks are textured with extracts from the photographs used during the photogrammetric process.

Each photograph contains more than one block. There is therefore room for some optimisation given that the loading of the photographs is quite CPU and memory intensive. Photographs are consequently chosen in order to minimize their memory load. However, the user can still choose to display another corresponding photograph of their choice for each block.

Figure 8 & 9 shows an archaeologist using the HIVE to interface with the 3D Shawbak Castle. Our user wears head tracked stereo shutter glasses in order to experience an immersive representation of the data. The user interfaces with the system via a tracked pointing device that allows the user to select between two modes: flight or block selection. Flight mode permits the user to fly around the castle by positioning and manoeuvring the pen in 3D space. We restrict the user from rolling the camera as this often complicates interaction. In block selection mode, the user can physically point at blocks on the display and select (highlight) those of interest. In Figure 8, the user is interested in interrogating the top left hand side of the arch and has consequently highlighted those blocks (shown in pink). Highlighted

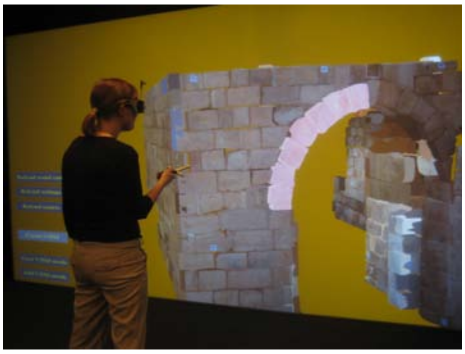


Figure 8: *Immersive visualization of part of the castle wall using a 6x2.5m^2 rear projected stereo wall.*

blocks can then be temporarily removed from the display, or, any block not highlighted can be temporarily removed. A menu system on the bottom left of the display permits the user to further interrogate highlighted blocks and extract further block information stored within the remote database. For example, stratigraphic information or volume of mortar used within the highlighted blocks. The user can also make accurate 3D measurements within the virtual environment.
While the user is navigating through the blocks and selecting them, requests are continuously sent to the database in order to visualize the stratigraphic units and how they are related to other units. The archaeological data which are displayed when a block is selected comes from the same data repository as the geometry, thus providing a better consistency of data. Our 3D immersive display is not limited to display environments such as the HIVE. The same software will run on a standard desktop PC or laptop (with or without stereo). In this instance, the user can interact with the system using a 2D mouse.

5. Conclusion and future work

The work presented here is the first result of a collaboration between three laboratories. It shows the feasibility of the program and will be used as a basis for future developments. Regarding architectural and archaeological analysis, three essential issues of this project are to be identified: the easy

Figure 9: *Stereo visualization of part of the castle wall.*

method for directly measuring structural elements. The I-MAGE process (Image processing and Measure Assisted by GEometrical primitive), is a useful device for the user and this time implemented in a specific module dedicated to the Shawbak survey. After having measured some points on the object-surface, the system is set up to let the user focus their attention on semantics, considering pictures one by one, while the system automatically completes 3D measurements.

The link between the scene representation in 3D and a database. This second feature allows combining a representation of the architecture itself to the database serving as a tool for the analysis of its units. The use of a three-dimensional model as a user-interface to the data allows linking the purely documentary data (references, observations made during the excavation, photographs) to a 3D representation of the object. This graphical expression of the object relies on geometrical data (position, orientation, dimensions) as well as "knowledge" of the object (theoretical shape, default values, relationships between diverse objects). The 3D model, produced by the system, shows the generic model of the object, defined by the archaeologists, and measured by photogrammetry and thereby becomes a relevant interface between the user and the collected data.

The third point is the use of architectural knowledge in order to perform the measurement process: knowledge is used to get a 3D model of each block by extrusion and data fusion. This approach allows inconsistency check and insures the database is always valid.
Finally our immersive visualization system permits the user to stand next to a photorealistic digital model of Shawbak Castle that has been created dynamically from a remote archaeological relational database. The user can also interrogate the database using predefined SQL queries. Future work will permit the user to construct any database query within the virtual environment and will involve the development of

novel techniques for the visualization of temporal and multivariate block data.

6. Acknowledgements

We would like to express our deep thanks to Serena Borchi, University of Florence, who gave precious help orienting all the photographs with photomodeler, Laure Lopez, MAP-GAMSAU laboratory, Elisa Pruno, University of Florence and Aurelie Baumel from *Institut National des Sciences Appliquées, Strasbourg* who made the survey of all the blocks with Arpenteur. Particular thanks also to Saverio Caponi, philosopher, University of Florence, who started an interesting discussion relating to the double point of view needed to see the building both as a collection of architectonic objects and also as a collection of atomic measured blocks. The first step of his work is published here: [Cap04].

References

[Cap04] CAPONI S.: *I mattoni dellâĂŹarco*, university of firenze ed. Department of Philosophy, Firenze, 2004.

[DDG01] DRAP P., DURAND A., GAILLARD G.: Simple photogrammetric methods with arpenteur. 3-d plotting and orthoimage generation : the i-mage process. In *CIPA 2001 International Symposium* (Potsdam, Germany, 2001), University P., (Ed.).

[DDS*05] DRAP P., DURAND A., SEINTURIER J., VANNINI G., NUCCIOTTI M.: Full xml documentation from photogrammetric survey to 3d visualization. the case study of shawbak castle in jodan. In *CIPA 2005 XX International Symposium* (Torino, 2005), ISPRS, (Ed.).

[DG00] DRAP P., GRUSSENMEYER P.: A digital photogrammetric workstation on the web. *ournal of Photogrammetry and Remote Sensing 55*, 1 (2000).

[DHVG00] DRAP P., HARTMANN-VIRNICH, GRUSSENMEYER P.: Photogrammetric stone-by-stone survey and archaeological knowledge, an application on the romanesque priory church notre-dame d'aleyrac. In *VAST - Euroconference on Virtual Archaeology* (Arezzo Italy, 2000).

[PCVH04] PANSIOT J., CHAPMAN P., VIANT W., HALKON P.: New perspectives on ancient landscapes: A case study of the foulness valley. In *The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage* (Belgium, 2004), pp. 251–260.

[VALJ02] VOTE E., ACEVEDO D., LAIDLAW D., JOUKOWSKY J.: Discovering petra: Archaeological analysis. *VR. IEEE Computer Graphics and Applications* (2002).

[VN03] VANNINI G., NUCCIOTTI: Fondazione e riuso dei luoghi forti nella transgiordania crociata. la messa a punto di un sistema territoriale di controllo della valle di petra. In *SAMI 2003 prodeedings* (Firenze, Insegna del Giglio, 2003), SAMI 2003.

[VNTD02] VANNINI G., NUCCIOTTI M., TONGHINI C., DESIDERI A. V.: Medieval petra. archaeology of crusader-ayyubid fortified settlements in trans-jordan. In *Civilisations of the Past, dialogue of the present: Italian Research mission in Jordan* (Amman, Jordan, 2002), AA. VV.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

WWW-Based Building Information System for "Domus Severiana" Palace at Palatine in Rome by Open Source Software

Katja Heine[1], Christiane Brasse[1], Ulrike Wulf[2]

[1]Brandenburg University of Technology at Cottbus, Germany
[2]Deutsches Archäologisches Institut Berlin

Abstract

The aim of this research project is the digital documentation of the "Domus Severiana" Palace at Palatine in Rome through the integration of geometrical 3D- and non-geometrical information. The great density of information requires an efficient concept of data storage and management and an adequate means of presentation of high quality. The decision was made to develop a www-based information system based mainly on open source software modules. Linux, Apache, MySQL and PHP as well as VRML-data format as open source components, were used for the information system; only the construction of the 3D-geometrical model was realized by means of commercial software. The 3D-model shows the different construction phases and depicts the reconstruction ideas at the respective different construction phases. Above all the 3D-model is the basis of joining the building information system to other sources of information, available in databases: the digital "Raumbuch", which contains all non-geometric information about the several rooms, archives of plans, drawings and photos, catalogues of the devices, brick stamps and constructions, as well as keys to literature and archives. The data storage, management and analysis are the central tasks of the project; realistic visualisation is secondary.

1. "Domus Severiana" – part of the impressive imperial palace on the Palatine

One of the buildings of antiquity that have shaped the appearance of the city of Rome to this day is the Roman imperial palace on the Palatine, which in conjunction with the Circus Maximus, forms an impressive ensemble. The well-preserved, monumental ruins reach a height of more than 48 meters and have always been visible. In 1998 as a part of preparations for restoration the Soprintendenza Archeologica di Roma commissioned the departments of Building History and Surveying of the Brandenburg University of Technology at Cottbus with the investigation of this section of the imperial palaces. Historical research into the construction of the area of the "Domus Severiana" was conducted from 1999 to 2003. In 2002 the project was expanded to include studies of the "Garden Stadium" and the remains beneath the monastery of San Bonaventura.

Figure 1: *Building complex of the "Domus Severiana" at Palatine in Rome.*

At the beginning of the documentation process eight years ago there existed only confusing plans of the "Domus Severiana". One reason for this is probably that only a few, difficult to- interpret fragments remained of the architecture that rose above the main floor. In addition, the highly complicated organization of rooms, extending over several stories, resembles a labyrinth, which at first sight reveals nothing about the rooms' former use or their relation to a particular period or group of buildings. The few existing plans of the "Domus Severiana" show the building complex, which is organized over six stories, projected onto a single plane without any differentiation of the different heights. Therefore the primary concern of the current documentation has been to represent the building according to its planes in six sections, which in turn can be clearly related to each other. Faced with the considerable height and complexity of the ruin and a surface area of 300,000 square meters, distributed in some cases on vast and poorly lit levels, this has not been an easy task. It could be achieved only through the use of the most modern surveying technology. The methods of recording and surveying the entire installation have been described in detail in [WR00], [Wul01].

Aside from the traditional tachymetry with reflector, tachymetry without reflector was applied as well as photogrammetry and traditional manual survey. All measured points are related to a coordinate system. The measured data underwent a computer-based revision and thus are available for further drawings, for instance, with the AutoCAD system or as 3-D coordinates for a three-dimensional mode.

On the basis of this documentation numerous individual observations on building details (such as joints, seams, various types of brickwork, foundation and vault techniques as well as surface and decorative details and extant brick stamps) were recorded in a room inventory ("Raumbuch"), which allows the systematic registration of all relevant details for the 186 rooms. In order to recapitulate the complex findings three-dimensionally and to more easily track changes about the sequence of building phases, a three-dimensional model of the entire documentation was made. It also serves as a departure point for reconstruction considerations.

2. Requirements for the Building Information System

The investigation of the building structure and the chronology of the building phases as well as the documentation of the investigation results required the installation of a digital Building Information System (BIS), that is, a 3D-Geographic Information System (GIS), which integrates geometrical data and non-geometrical information. The main components of the BIS are the 3D geometrical model of the building complex set up by AutoCAD software and databases for the special non-geometrical information like "Raumbuch". Although there are many commercial GIS-software products available, for real 3D application only a few are offered. The requirements to the BIS for the "Domus Severiana" are as follows:

- Integration of 3D geometrical data
- Integration of very complex non-geometrical information with a high density
- Centralized data management
- Decentralized access to the information from any place worldwide
- Efficient publication of research results
- Easy to operate software for non-specialists
- No license cost for data reader / viewer

From an analysis of several products it was realised that none of them was able to fulfill all of the requirements of the BIS and thus an independent solution was necessary to develop. It was decided to realize this on the base of open source software modules. Apart from the advantage that open source software has no license costs, the use of open source technology because of the open source code enables an optimal adaptation to user requirements. In addition, contrary to many proprietary programs open source software facilitates a standardized data exchange because of the use of standardized protocols and formats. High stability and security are characteristic of most open source software products, especially those with a wide distribution, due to intensive testing and development by the community of software developers. The decision for an open source solution often involves spending more time for installation and adaptation, but it may be a more sustainable solution than the alternatives.

For another research project of the participating institutions, the investigation of urban history of Baalbek / Lebanon, database modules by open source technology already were developed [HLFG05]. It made sense to use these database modules and to supply special databases like "Raumbuch" for the room inventory to the database system. The advantage of this solution is that the main part of the database system can be used by both projects.

3. Structure of the Building Information System

The complete open source LAMP-System was used for the implementation of the database system. LAMP means that the system software is LINUX and the server software is Apache. MySQL (`http://www.mysql.com`) is used for the databases and the scripting language PHP (`http://www.php.net`) accomplishes the communication between www-client and the database system server (see fig. 2). LAMP is a very common open source technology, which is used worldwide even by large companies. Geometrical information is depicted by VRML (Virtual Reality Modelling Language), and the usage of X3D data format is planned. Both are standardized data exchange formats for 3D graphical information on the World Wide Web. The user needs only an Internet browser and a free available viewer like CORTONA VRML client (`http://www.parallegraphics.com`) or Octaga player (`http://www.octaga.com`). All these open source components are widely used and well accepted.

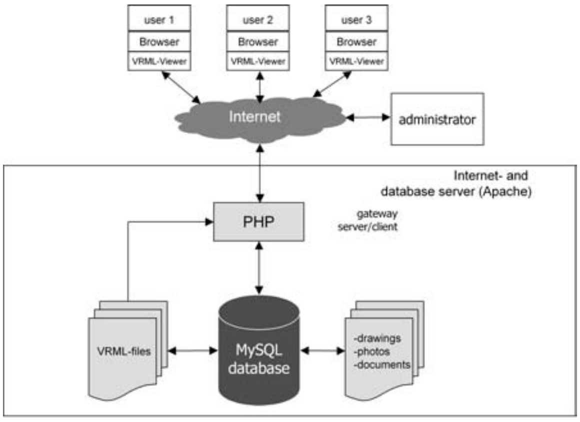


Figure 2: *Technology of the www-based building information system.*

3.1. Design and Functionality of the database system

The database system is set up modularly (see fig. 3). That means it consists of several databases. Some of these modules have a basic character such as the bibliographical database and databases for photos and drawings. Others, like "Raumbuch" for the room inventory, as well as databases for architectural fragments and for archaeological finds have specific contents. All information relevant to the building has to be stored in the database system in a consistent way. In addition the system structure has to allow for complex queries using relevant search tools. Databases for building research and archaeological purposes as a rule are characterized by a high number of attributes and a high degree of complexity. The conceptual phase of the implementation is very important and requires an intensive cooperation between users like architects or archaeologists and computer scientists.

Figure 3: *Modules of the information system.*

3.2. The concept of the database module for room inventory

For the information system of "Domus Severiana" three specific modules were developed. The database of room inventory, the database of architectural fragments, and the database for 3D- objects. The database module "Raumbuch" for the room inventory is of central importance. There are already good examples of architectural-fragments databases such as the database "Arachne" of the "Forschungsarchiv für Antike Plastik" of the University of Cologne (`http://www.archaeologie.uni-koeln.de`); thus the structure of "Arachne" database was used for the "Domus Severiana" project as well. The third special module of the system – the database of three-dimensional objects, consists of many data sets but with a simple structure, comparable to the structure of the basic modules like the photo database, so the realisation was quite easy.

During the campaign on the Palatine, an analog catalogue, organized according to the six main levels, has been produced. The use of the "Raumbuch" was necessary to collect all information which could not be represented by plans. For every room there is a data sheet that consists of textual information and all views as sketches, on which all construction details are registered, along with important measurements and information about joints, building periods, earlier fittings or furnishings, dating criteria, and peculiarities in technique, brick stamps, etc. This analog catalogue cannot be used for analysis, of course, because it requires an extensive time to search and find various information. A more practical instrument is the digital "Raumbuch" database which allows one to archive and work with a large mass of non-geometric information and to use selective queries.

The concept of a very flexible and comprehensive structure is a basic condition, but it requires a great deal of time and effort. The most important requirement of a database is the well-defined and consistent assignment of information to the different objects. It quickly became clear, that the structure of the analog "Raumbuch", which assigns all information to the rooms, would have to be more detailed. The digital database not only stores all information about the rooms and special attributes of structural elements, but it must also represent the topology of geometrical constructive objects. Topology describes the invariant connections and relationships between objects for instance, which wall elements form the boundary of a particular room. Indeed the room is the superior class. However, only with the surrounding constructive elements does a room become visible. These constructive elements are classified in different groups; there are horizontal elements – (floors, slabs or vaults), vertical elements (walls) and other elements (sewers, pillars and columns). These elements belong to the subordinate object classes and can later be linked to the several objects of the three-dimensional model. Nevertheless these elements contain more specialized constructive information such as the openings (doors, windows or niches) or the very important brick stamps. These objects are collected in their respective classes and are linked with the constructive elements. On one hand constructive elements with linked special elements can

Figure 4: *Data sheet of a room with linked information (marked by the blue line).*

be assigned to the different rooms, and on the other hand the detailed structure enables the assignment of a certain wall to several rooms (see fig. 4).

Another class is composed of the decoration of a room; such a class may contain, for example, mosaics or plaster, which are connected with constructive elements. Because a wall can have a different decoration in different rooms, the decoration is assigned with the room and not with the constructive elements.

The detailed structure of different classes enables a hierarchical assignment. With the inherited attributes all information about a room can be indicated or selected, without having to include an object into various classes. To simplify handling, every class consists of a similar structure with similar terms to describe the objects. For example the terms of typology, interpretation, or state of preservation are used in all classes, of course with variable content.

Information of the basic modules can be linked with the objects of the "Raumbuch" in the same way. The previously described sketches of the analog "Raumbuch" are collected in the plan database and can be linked with different rooms. When that is done a system is produced which includes all information of the analog "Raumbuch" and the other archives in a well-defined way. This complex database is a useful instrument for working, searching, and analysing the information of the "Domus Severiana".

The hierarchic structure of the database "Raumbuch" (see fig. 5) is equivalent to the hierarchic construction of a common building; hence this module can be used for other projects too, without any changes.

3.3. Modelling geometry

In order to display the complex information contained in the different levels of the "Domus Severiana" a three-dimensional model was created. To produce the CAD model, measurements and the scanned set of plans are integrated and referenced to Rome's city map. Measurements and plans form the geometrical basis for abstract representation gained from simple 3D cubes. The preciseness of the CAD-model cannot correspond closely to the criteria used for the documentation of the building; the amount of work such precision would require would not be justifiable. Although the abstract model shows the irregularities and particularities of the building, it does not give a detailed rendering. The creation of the CAD model has been described in more detail in [RW02]. The 3D-model facilitates a better navigation through the installation and offers a better visualisation of the existing building (see fig. 6). In addition, the model provides for the analysis of complex structures and can be used as a basis for reconstruction ideas of the different building periods. Ideas related to load-bearing construction, construction sequences, and use of and access to different units, can be checked quickly and can be represented three-dimensionally. Irregularities in the building structure

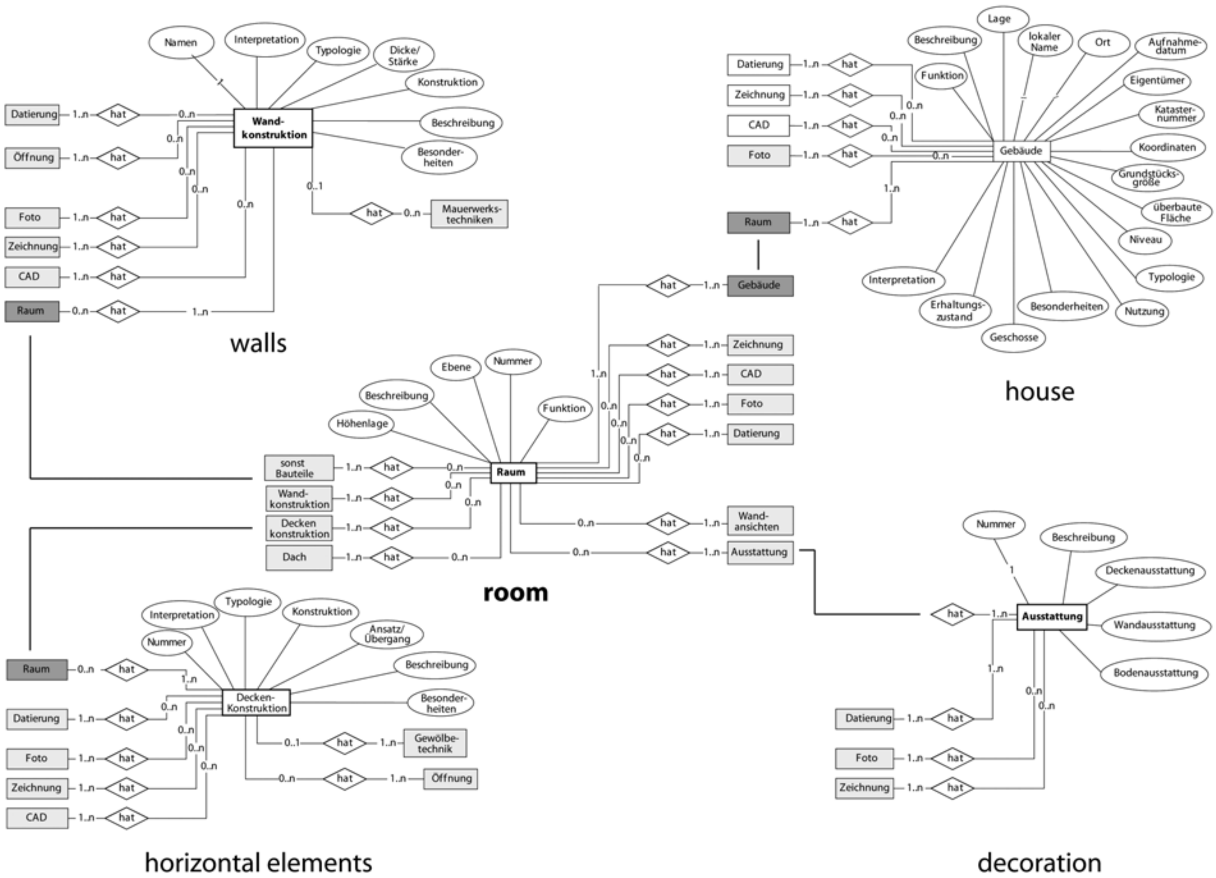


Figure 5: *Concept of the database "Raumbuch" for the room inventory.*

(e.g. the position of load- bearing walls in the different levels) become clear while creating the model.

3.4. Visualization of geometrical data

The 3D-model is composed of individual room elements. So it represents the same topological structure as the "Raumbuch" database. Corresponding elements of both modules can be linked. For the visualization of the geometrical model on the Internet, the AutoCAD data are transformed into VRML data format. The export of dwg- or dxf-files into VRML-files is possible because almost every CAD- or 3Dmodelling program includes an export function to VRML file. These VRML-files can be displayed by a 3D-Viewer, which includes the necessary, simple functions of navigation. The Viewer also offers the options of interior perspectives or virtual walks through the complex. The technical advantage is the very small size of the files. The whole model of "Domus Severiana" consists of more than two thousand objects but the file size is only 1.7 MB. Such a minimal size guarantees fast processing on the Web. The second advantage is the open code of the files, which allows for the storage of information of the several 3D-objects in the database.

3.5. Data analysis and queries

The connection between non-geometrical data stored in the databases and the 3D geometrical model has to be realized in such a way that two tasks can be performed: One, it must be possible to get information from the database about building objects such as walls while one is navigating through the 3D model. Two, the result of a SQL-query done in the database, for instance a wall with special attributes, has to be displayed. It is indispensable to utilize object keys, which are identical in the database and in the geometrical model. The geometry of every object, i.e. every wall in VRML format is stored in the database. The description of geometry and appearance of the particular objects were filtered out from the general building model by a parser. With the key the 3D objects and the objects of "Raumbuch" can be linked. If a query results in a set of objects, a temporary VRML file set up from the geometrical description of the individual objects is compiled. The appearance of the objects depends on the dating

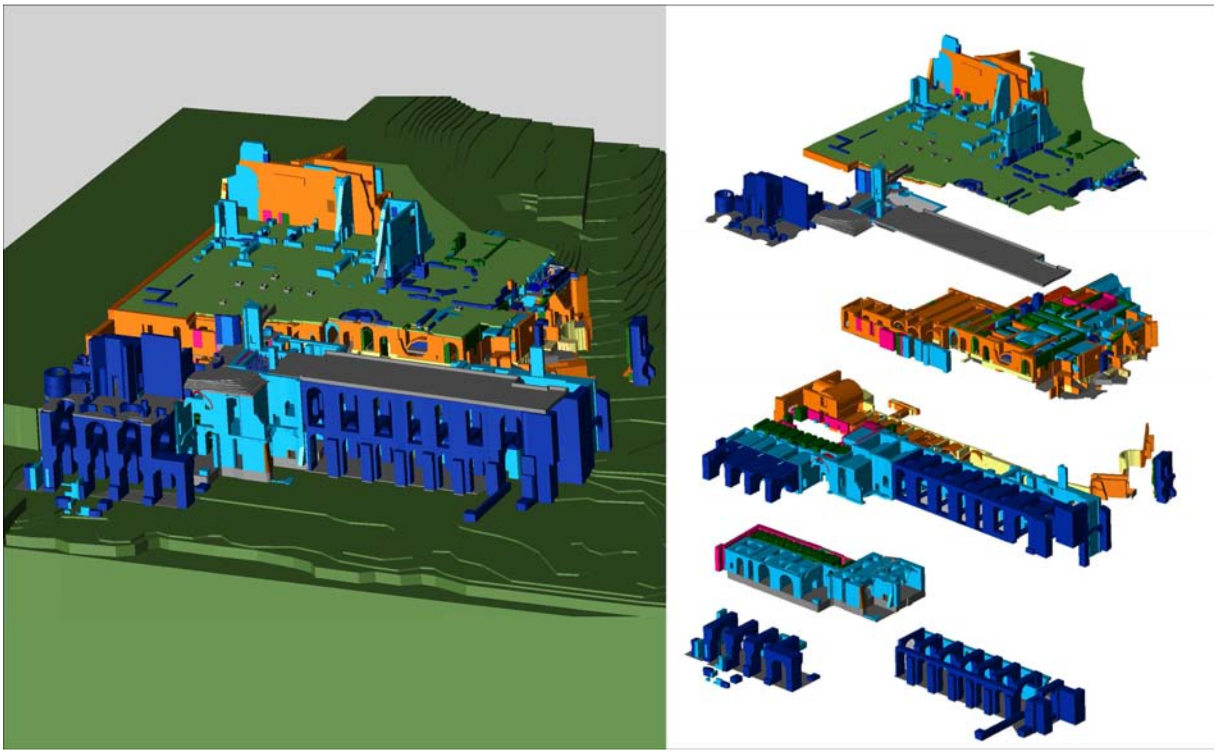


Figure 6: *3D geometrical model with several levels of the building complex.*

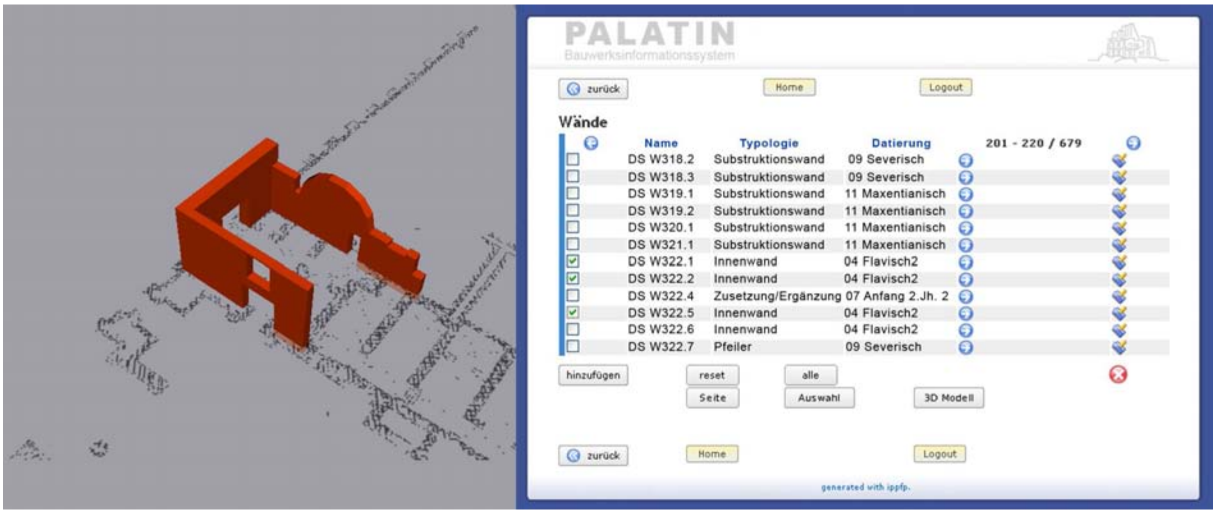


Figure 7: *Result of a query, visualized by a temporary VRML file.*

and thus on the building period the object belongs to. For example one can search in the database for all walls with brick stamps in the Flavian phase, and then all these wall objects are visualized in the model. Additional 2D ground planes are included into the VRML files for a better overall view and navigation (see fig. 7). The simple code of VRML enables the control of object appearance in a VRML-file depending on the value of a certain attribute such as the building period. If the building period is changed, the colour of the 3D-object is changed automatically. Thus new assumptions can be checked easily by the model.

The inverse direction from the VRML model to the database is to be solved by including hyperlinks from the geometrical object to the corresponding data set by a passing parameter in the form of an object key. The integration of hyperlinks is realized in all typical VRML-viewers. When one clicks on an object with the hyperlink icon, a program routine detects which object was selected. The key number of this object is included as a parameter on the http-address for the database request.

The function of visualizing the layer-structure of AutoCAD is an important function used to show and hide objects. This possibility doesn't exist in typical 3D-Viewers. Programming of a special function is necessary to reach better functionality for navigating through the model.

4. Research results of building history received from BIS

The 3D model, the analysis of the information stored in the database, and the plans of the building phases make it possible to look at individual phases one by one: the first phase, the pre- Flavian structures (middle of the 1st c. A.D) cannot be interpreted with certainty. The second phase, the first Flavian phase, brought to light one of the most surprising results of this building analysis: as early as the end of the first century A.D. the Flavian imperial palace extended far beyond the so far known area, the delineations of the "Domus Augustana" and the adjacent "Garden Stadium". A substructure formed an artificially created huge platform for the main story with a row of viewing rooms, which opened with tall windows to a huge water basin. Towards the southeast, facing the Circus Maximus, this area terminated with a colonnade that may have opened with a column or pillar arrangement that allowed for a view to the Circus and the surrounding landscape.

Except for an addition, also from the Flavian time, and an intermediate phase that can be dated to the 2nd century, no larger building activities can be detected as having taken place until Severan times. A large-scale destruction, which most probably occurred in connection with a huge fire that has been documented for the winter of 190/191 A. D., must have preceded the Severan changes of use and additions. The main story must have suffered so much during the fire that it was almost completely rebuilt, though apparently the ground plan was retained. Septimus Severus not only renovated this part of the palace but also extended it in order to create space for a new bathing complex high above the Circus Maximus.

The end of building activities in the area of the "Domus Severiana" is marked by a further generous extension of the Severan bathing complex. Once again, the platform was considerably extended. The springing of vaulted arches shows that the installation was much more extensive than today's visible remnants. The bathing rooms occupied a large area of the platform and the former viewing rooms were incorporated into the architecture of the bathing complex. Unfortunately nothing can be said about the individual look of the additional rooms because today nearly all installations are lost.

The documentation and architectural-historical investigations of an extensive and complex installation such as the "Domus Severiana" couldn't be realized within the justifiable limits of time and financial expenditure without the application of modern technology. Also it is necessary to use innovative, new methods and media for evaluating and presenting the results of the detailed investigations as they were carried out with thousands of individual findings, which can be difficult to track, especially in such a three-dimensional complex and confusing area.

5. Prospects

The information system is a practical and effective instrument to store and analyse non-geometric information and to link this information to the 3D geometry of the great building complex of the "Domus Severiana". Due to the Internet, this information is available to users across the globe. The use of open source software and formats guarantees the access to this information today and in the future without limitations.

The concept of the BIS is not limited to the "Domus Severiana" project. Rather it shall be used also for other areas and buildings on the Palatine like "Domus Augustana" and the "Garden Stadium". The structure of the databases of the Building Information System and the object classes and attributes included are conceptualised in a generalised form, so that the application to other buildings can be easily realized.

Furthermore the intention is to set up a basic information system for building historical research projects in general. By including database modules and solutions developed for the already mentioned Baalbek-project, e.g. for visualisation of 2D geometries by a MapServer, a multi-pur pose information system applicable to any building historical research projects can be developed. Two basic problems have to be solved: One an ontology of building research data for creating a semantic standard has to be worked out, and two existing standards for modelling geometry such as standards of Open Geospatial Consortium (OGC) and ISO TC 211 as well World Wide Web Consortium (W3C) standards have to be strictly considered. The concept of a basic information system can be successful and well accepted only if these preconditions are met.

A basic information system for historical building research would be a helpful support of research projects in building history and at the same time it would offer an efficient tool of publication, exchange, and archiving of research data and results.

Acknowledgement

We are grateful to the Soprintenza Archeologica di Roma, division Palatin, represented by I. Iacopi, M. A. Tomei and its former director A. LaRegina. The project receives many forms of support from the German Archaeological Institute (division Rome). Fritz Thyssen Foundation and Deutsche Forschungsgemeinschaft DFG support the project financially. Gerry Neumann and Dexu Zhao programmed data base structure and software tools.

References

[HLFG05] HENZE F., LEHMANN H., FISCHER-GENZ B.: Development of an internet-based information system for archaeological research and studies on urban history in Baalbek/Lebanon. In *XX. CIPA-Symposium, Turin* (2005), pp. 849–854.

[RW02] RIEDEL A., WEFERLING U.: From pencil drawing to computer model: a method combining strategy for the documentation of the "Domus Severiana" at the Palatine in Rom. *Surveying and documentation of historic buildings-monuments-sites* (2002), 132–139. (Proc. XVIII int. Symposium CIPA 2001, ed. Albertz J.).

[WR00] WEFERLING U., RITTER B.: Bauaufnahme der "Domus Severiana" auf dem Palatin in Rom – eine interdisziplinäre Aufgabe für Architekten und Geodäten. *Mitteilungen des Deutschen Archäologischen Institutes, Römische Abteilung 107* (2000), 299–310.

[Wul01] WULF U.: Mit welcher Methode sollen wir aufnehmen? Kombinierter Einsatz von Aufnahmemethoden am Beispiel der "Domus Severiana" auf dem Palatin in Rom. *Von Handaufmaß bis High Tech* (2001), 153–164. (Proc. Von Handaufmaß bis High Tech – Aufnahmeverfahren in der historischen Bauforschung, Cottbus 2000, ed. Weferling U., Heine K., Wulf U.).

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Interactive Mobile Assistants for Added-value Cultural Contents

M.T. Linaza[1], A. García[1], A. Susperregui[2] and C. Lamsfus[3]

[1]Visual Communication Technologies VICOMTech, Spain
[2]Polytechnical University College, University of the Basque Country, Spain
[3]De Valck Consultants, Belgium

Abstract

Multimedia technologies provide new opportunities for museums to enhance their visitors' experience. However, its use brings new challenges for presentation preparation, among which are how to enrich the visit while not diverting the visitors' attention from the actual objects in the museum, which should remain the focus of the visit; and how to provide a rich information space suitable for a wide variety of visitors. These challenges need to be addressed during planning and preparation of information presentations for mobile, multimedia museum visitors' guides.
This paper describes the design and implementation of the AMICo prototype to an exhibition room for visitors who are equipped with wirelessly connected handheld devices. The prototype has been implemented on an exhibition showing a set of architectural scale models from the famous Spanish architect Rafael Moneo hold in the Kubo exhibition centre of the Kutxa Foundation in Donostia-San Sebastian during September and October 2005. The architectural scale models were augmented with information in the form of multimedia content. Users were able to access those contents in a personalized way when in the proximity of the artwork.

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

1. Introduction

Museums and exhibitions do not communicate. These places are often only a collection of objects, standing deaf in front of the visitors. In many cases, objects are enhanced with textual descriptions, usually too short or too long to be useful for the visitor. In the last decade, progress in multimedia has allowed for new, experimental forms of communication in public spaces using computer technologies.

Implementations range from simply using standard PCs with multimedia applications to large theatres that immerse users in virtual worlds or reproduce and display 3D models of masterpieces. Content generators often just transfer the available technology to traditional museum schemes, without paying much attention to the ways they expect users to interact with the system or to the cognitive and aesthetic factors involved.

During the last decades, a large number of projects related to the application of innovative technologies to Cultural Heritage has been implemented. Several European projects have assessed some of the most outstanding technologies: 3D digitalization and scanning techniques have been used for the reconstruction of historical objects (3D-MURALE); Virtual and Augmented Reality technologies allow new interaction ways for users and experts (ARCHEOGUIDE, RENAISSANCE, CHARISMATIC); and mobile devices and multimodal interfaces provide intuitive and personalized access to scientific information from museums and archaeological sites.

Concerning multimedia mobile guides, they must support important personalization of the content owned by a cultural institution in order to provide users with an enjoyable visit in accordance with his/her background. At the same time, a guide for a museum or exhibition room should encourage learning and personal enrichment. Therefore, information should be displayed taking into account the physical location of the visitor and the position of the artworks in their natural environment. The overall experience can be optimized connecting the information of the exhibition and the presentation to the visitor in a coherent way depending on the location.

This paper describes the design and implementation of a prototype for visitors to an exhibition equipped with wirelessly connected handheld devices. Section II of the paper presents a brief state of the art about current applications concerning electronic guides in cultural environments. Within Section III, the project is described, including the main requirements and objectives, the design and implementation of the AMICo prototype and the validation scenario.

Section IV presents the modules of the prototype: the relational database for the existing information about the exhibition, the author and the artworks; the Authoring Tool implemented for the generation of multimedia personalized added-value contents; the tracking system based on wireless technologies; the visualization device and the tools for the assessment of the behaviour of the visitor. Finally, some conclusions are provided.

2. State of the art

The growing availability of portable devices with increasing computational and interactive resources in terms of power and capacity has raised research issues regarding the design of user interfaces for applications with such devices. Currently, museums and cultural institutions are becoming an attractive application area for such interactive mobile devices.

Museum visitors can be assisted in several ways. One possibility is the use of Audio Tours: visitors can use a sort of large telephone receivers. They can select the work of interest typing its code through a numeric keypad. Audio Tours are useful precursors of electronic guides, although they have limited visual channel.

Museums have used portable electronic guides or eBooks since the 50s. Although they have already had a long history, these guides are still raising research issues regarding the use of portable computers. Electronic books allow displaying information and disseminating content during the interaction with different objects of an exhibition. Currently, multimedia guides display static images or pre-recorded videos describing the artworks of an exhibition.

One of the first fully operating systems in a museum setting was implemented at the Marble Museum of Carrara [CP04]. The managers of the museum decided to provide their visitors with additional information to that available in traditional formats such as leaflets or audioguides. The system stores the information locally in the memory of the PDA, uses a map to guide the visitors around the museum and presents content of different abstraction levels (room, section and exhibit).

The Sotto Voce system implemented by Xerox analysed the use of mobile technologies in exhibitions [AGH*02]. The project was validated in different exhibition rooms of Filoli, an historical house in Woodside, California. Each visitor carried a Compaq iPAQ device with a tactile screen, a wireless card with Internet protocols and a single-earphone. The evaluation of the prototype has been positive, because a balance expenditure of the time of the visit between the electronic books and the real artworks was achieved.

Zancanaro, Alfaro and Stock [ZAS03] have implemented kinematics techniques for the visualization of some details of the frescos at Torre Aquila, a tower of the Castle in Trento. The multimedia guide, implemented with Macromedia Flash on a PDA, detects the position of the visitor by means of infrared emitters placed in front of each panel. The multimedia presentation is composed of an audio commentary accompanied by a sequence of images that appears on the PDA display and helps the visitor quickly identifying the details of the fresco mentioned in the commentary.

Headquartered at the Exploratorium in San Francisco, the Electronic Guidebook project [RS02] is a study of visitors for the evaluation of the use of wireless network technologies in order to "bridge the physical and virtual worlds". Visitors to the science museum were equipped with wirelessly connected handheld devices. The museum exhibits were augmented with information and services based on Web technology, so that users could access the proper contents when in proximity of the exhibits.

Sponsored by Bloomberg, the Tate Modern's Multimedia Tour Pilot was open to the public from July to September 2002 [Wil04]. Developed by Antenna Audio in collaboration with Tate Modern, the multimedia tour was a 45-minute tour of the Still Life/Object/ Real Life galleries, in which visitors could experience audio, video, still images and a variety of interactive applications on handheld iPAQ computers loaned by HP. The content of the multimedia tour was delivered to the visitor through the wireless network of the museum, using location-based technologies.

With the technological support of Sun Microsystems, the J.Paul Getty museum has developed the GettyGuide project in order to supplement its network of information kiosks. The project is based on Palms as the mobile platform, and uses Flash to display the multimedia content received from the museum central server via its wireless network. Moreover, visitors can search for artworks by queries using several criteria such as artist or genre.

Finally, the main goal of the MoMo system is the design of a user interface for multimedia applications on small devices such as PocketPCs [JEMC05]. The key issue faced by this system is the way information must be displayed so that it is assimilated by the user in the most adequate way. The project introduces a new type of browsing graphical control in the framework to explore collections of arbitrary related information.

For further information, the reader is referred to [RTA05], where a detailed review of mobile applications in museum environments is provided, focusing on the notion of context.

3. Description of the AMICo project

The AMICo project analyses the use of wireless technologies in the implementation of interactive mobile assistants based on Virtual Reality and storytelling technologies. The availability of portable devices with wireless capabilities has a great potential in the Cultural Heritage application domain. Using the wireless networks, mobile devices can detect the position of the visitor and stream multimedia content in real time.

The project relies on a high-bandwidth wireless infrastructure to download information about artworks as visitors walk through the exhibition. This enables supporting them with interactive services and highly dynamic information. In addition, the same network infrastructure provides location information to the system.

The AMICo prototype focuses on three simultaneous areas of research: information technology infrastructure (network components delivering the information and tracking the user); human computer interface issues (usability, interfaces); and content development (design, formatting). The project has been designed as a proof of concept study to explore potential areas for further research and development.

Finally, one of the main objectives of this project was the evaluation of the behaviour of the visitor when new technologies are introduced in cultural and artistic environments. A deeper knowledge about human behaviour in cultural institutions in relation to concrete technologies approaches has been generated. The results of the usability study will be published in further scientific papers.

3.1. Requirements of the prototype

The AMICo prototype includes two different components: the mobile device that is carried by the user and the server that provides multimedia content. The multimedia mobile device is the interaction interface between the user and the system. Its main requirements include wireless connectivity with IEEE 802.11 networks, portability (weight) and support for multimedia presentations (size, colours, resolution).

The server is responsible for collecting and processing queries from the mobile devices; determining the location of the mobile guides on the basis of the data coming from the tracking system; managing and rendering the personalized added-value content; and storing the results of the evaluation process.

Wireless communication technologies are a key issue within the project due to the different available mobile devices and solutions (colours, size and weight, processing capacity). Another important aspect is the ability of the mobile

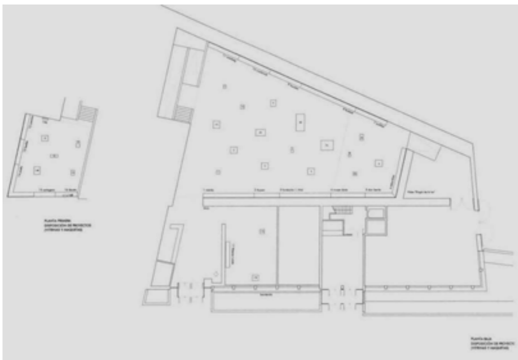


Figure 1: *Layout of the 17 architectural scale models.*

devices to be tracked by an accurate system for indoor applications (RFID, WLAN, Bluetooth). Moreover, the wireless network should be designed in order to support the simultaneous performance of several mobile devices with the proper bandwidth.

Regarding the contents available for the user, while traditional golden rules for interface design are a basic reference in the field of Human-Computer Interaction, there are no equivalent guidelines for mobile devices. The main requirements that have to be fulfilled by those contents are:

- Functionality. Graphical user interfaces should not require long learning curves, avoiding approximations that may not be intuitive for the user. By doing so, we contribute to relieve efforts of the users when interacting with the prototype and make their experience more attractive.
- Simplicity. The information provided to the user must be multimedia in nature. Multimedia information is very familiar to the human perception and, consequently, very attractive. It should be taken into account that although there could be large amounts of information to be potentially accessed, this process should be transparent to the user.
- Flexibility. The solution should be adaptable to the range of domains in which it might potentially be used: museums, monuments or exhibition rooms.

3.2. The AMICo prototype

The AMICo prototype has been implemented on an exhibition showing a set of architectural scale models from the famous Spanish architect Rafael Moneo hold in the Kubo exhibition centre of the Kutxa Foundation in Donostia-San Sebastian during September and October 2005. As a curiosity, it must be mentioned that the exhibition centre itself is placed inside one of the buildings designed by Moneo. The exhibition included 17 architecture scale models distributed in three exhibition rooms as shown in Figure 1.

Visitors carry the prototype with them as they walk

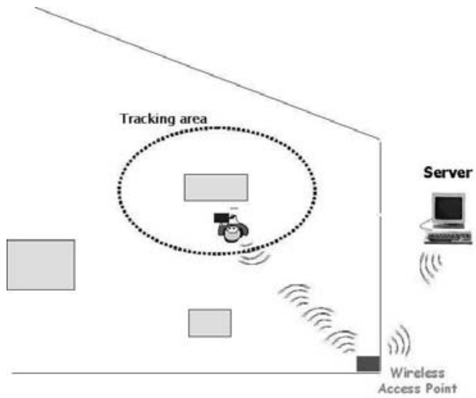


Figure 2: *Performance of the AMICo prototype.*

through the exhibition. When the mobile multimedia guide detects that a visitor is in the nearby of a specific scale model, general information about the scale model is displayed, including its main characteristics, texts and additional images (Figure 2). Furthermore, the system provides further information about specific topics, such as architectural details or materials of the real building.

For example, one of the most outstanding scale models was the future Church of Riberas de Loyola, a new district that is being built in San Sebastian. When approaching the scale model, visitors could visualize multimedia content on their handheld guidebook, so that they could perceive the future visual impact of the Church within its environment.

The multimedia presentation consists of an audio commentary combined with a sequence of images that are streamed and rendered in real time in order to help visitors quickly identifying the scale model and the real building. At any given moment, the user is free to pause or even stop the presentation by clicking on the appropriate control panel button.

As it has been mentioned, one of the aims of the project has been the usability study of interactive mobile assistant devices in real exhibitions. Therefore, the visualization application traces and records the selections of the visitors, the concepts that have been visualized, the time spent on each information piece and the concepts required for more information. At the end of the visit, a short questionnaire is provided to the user in order to gain information about further improvements of the system.

4. Technological components

4.1. Overview of the system

The prototype includes five main components: storage of the existing multimedia information about the buildings and the author using relational databases; generation, storage and management of added-value content with an Authoring Tool; tracking of the user based on wireless technologies; visualization of the multimedia content and evaluation of the behaviour of the visitor.

4.2. Storage of existing information

Within six months, multimedia information about the scale models has been collected, including technical and aesthetical aspects of the real buildings, biographic data about Moneo, audiovisual material about the building process and additional contents.

We have analysed the different existing approaches for the storage of the information, including relational databases, where all the available data is organised in value tables; object-oriented databases; and information repositories, where the information is not structured due to its high dynamism.

After a deep study of the requirements of the system, a relational database has been implemented in order to take advantage of the associated Data Base Management System (DBMS) that helps managing the database. A DBMS is a group of programmes, procedures and languages that allows working on a database without accessing the programming language.

MySQL has been selected to implement the relational database. The final structure of the database has been compiled from the requirements and the proposed schema from the content generators. Main value tables include data about Rafael Moneo (complete name, place and date of birth), his academic studies (undergraduate studies, postgraduate, masters, PhD) and his outstanding buildings (name and place, objective of the building, technical and aesthetical aspects, materials, illumination).

4.3. Generation, storage and management of personalized added-value contents

The content generator uses the available information to generate the personalized added-value contents that are presented during the exhibition using an Authoring Tool. The generation process using the Tool consists of two main phases: the definition of concepts or contents, and the implementation of the associated web pages for each concept. The AMICo concepts are related to the architectural scale models of the exhibition.

Concepts are only created once and can be classified into two main categories: general and specific. The former includes generic information about the content. On the other hand, the latter is used when content generators want to offer specific information about some pieces of the content. When visualizing the information, the visitor is allowed to choose between both possibilities.

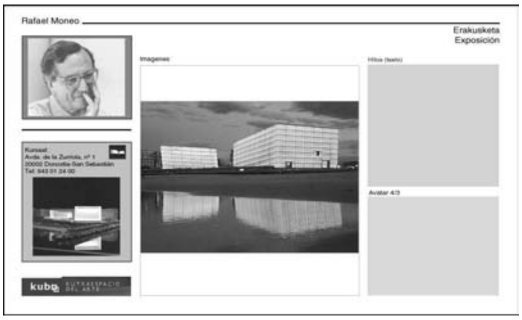


Figure 3: *Example of the template for the content of a scale model.*

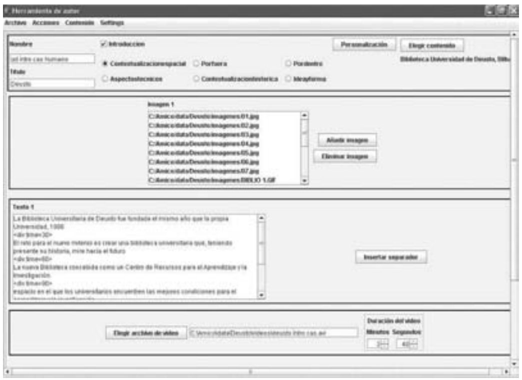


Figure 4: *Generation of added-value content using the AMICo Authoring Tool.*

Firstly, the appearance of the output is selected on the basis of some previously defined templates (Figure 3). These templates are dynamically created, allowing changing the distribution of the different types of content for visualization on the final web site. For instance, one template can include a field for text information and two fields for graphical information. Content generators determine the final choice of the template and the content.

Each content includes different types of information: images, texts, videos and audio files. Most of those types can be stored in sequences. This means, for example, that the content generator can specify a sequence of images to appear synchronized with the spoken text.

As content is personalized to the visitors, the same information must be presented to different profiles of users. Therefore, a profile is attached to the content during the generation process with the Authoring Tool. Although each concept includes different pieces of information, each of them is classified according to the user profile.

Once the added-value content has been created, it is stored in a second database for personalized contents, independent from the information database. This second database is owned by the Kubo exhibition centre as well as the added-value multimedia content that is visualized on the mobile device.

The second database has been filled by the content generators. The main advantage of storing the content in a database is that the output can have any format. Using an intermediate application, the information could be exported to different formats (a web site, audio). Another advantage is the possibility of adapting the content to different visualization devices.

The output of the storage and management system has a web site format. All the scale models have a main page whose URL is stored in the server, as well as several subpages. When the main page gets downloaded by the browser once the tracking system has detected the position, the visitor can click on some parts of the image to get further information about some of the related contents.

4.4. Tracking of the user

Location awareness is seen as a key issue within the AMICo project. Information about their position is not interesting for the users in most of the cases. However, if a system is aware of the current location of the user, it can infer what the most interesting contents are.

Within the project, we have selected the Ekahau software, a cost-effective solution for indoor location systems based on a Wi-Fi based location estimation system using the existing IEEE 802.11 network infrastructure. The selected approach can provide meter-level accuracy, which was assumed as acceptable for the AMICo validation scenario.

4.4.1. Selection of the tracking system

We have taken advantage of the high features of the visualization device, which is sensible to the use of wireless connections, in order to track the position of the user by means of the Ekahau software. The software including three components (Ekahau Client, Ekahau Positioning Engine (EPE) and Ekahau Manager) allows implementing location based applications for mobile devices working over a IEEE 802.11 network.

One of the main advantages of the Ekahau software is that there is no additional need for specific Access Points, because the existing ones can be used. The Kubo exhibition centre is composed of three exposition rooms, with a total area of 739.6 m2 distributed among three different exhibition rooms of 489, 146.6 and 104 m2 each. In order to cover the mentioned area, five Access Points have been installed, three on the main room and one on each of the remaining rooms (Figure 5).

The selection of the software was mainly based on the existing successful application scenarios. Even through the learning curve is quite complex, the installation process was quite simple.

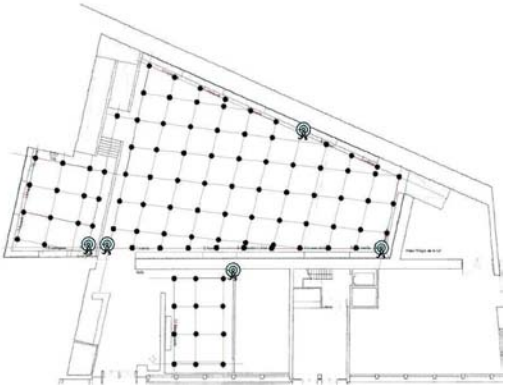


Figure 5: *Distribution of the Access Points.*

Figure 6: *Graphical interface of the implemented YAX application.*

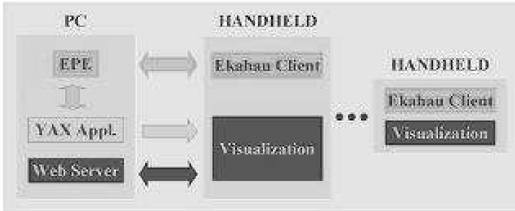


Figure 7: *Communication among the components of the tracking system.*

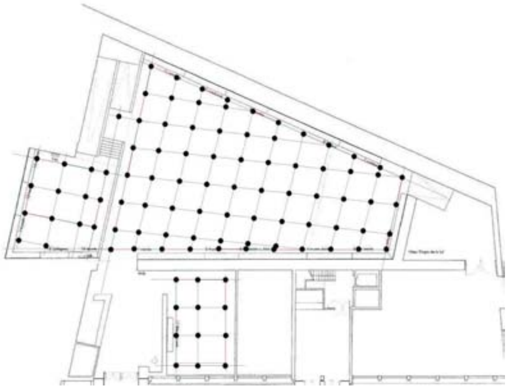


Figure 8: *AMICo and the grid of the WLAN tracking.*

4.4.2. Detection of the scale models

The Ekahau Positioning Engine (EPE) includes a standalone application with two different ways for detecting the position of the user: the Ekahau SDK for Java developers and the YAX protocol for any programming language supporting TCP sockets.

The YAX protocol has been implemented using the C++ object-oriented language, as the application for position tracking is independent from the content-supply application developed in Java. In addition, the YAX API has been completed with a plain graphic interface so as to help the personnel from the exhibition centre with the position tracking task.

The implemented YAK application allows selecting the desired scale model from the database, its position in the room, and the area around the artwork where the user should receive the information as shown in Figure 6.

Once all these data have been inserted, the program automatically detects which of the visitors should be provided with the proper information.

4.4.3. Performance of the tracking system

For a robust tracking, several aspects must be fulfilled: the EPE has to be installed in a computer, the Ekahau client should run on each handheld and a Wi-Fi net equipped with enough Access Points must be available. Figure 7 shows the modules of the prototype and the communication flow between them.

Two different phases are required for a correct calibration of the tracking system. Phase I can be called the offline training phase, in which we have performed a site survey by measuring the received signal strength indicator from different Access Points (APs) at some fixed sampled points of the exhibition room. Figure 8 shows the selected grid for the exhibition room. The measurements are recorded onto a radio map that depicts the strength indicator values of APs at different sampled points.

Phase II is known as the online estimation phase, in which the location of the scale models is calculated in real time by matching sampled points on the radio map with the closest strength indicator values to the target.

Figure 9: *The OQO as the visualization device.*

Figure 10: *Use of the prototype within the exhibition.*

Once the calibration process has been successfully achieved, the system performs as follows. The EPE is constantly tracking the position of the handhelds via the wireless LAN. The tracking data is retrieved by the YAX application, which calculates if users are inside the areas where information should be provided. If so, the identifier of the architectural scale model is sent to the Java application running on the handheld, which in turn asks the web server for the information relative to the mentioned identifier.

4.5. Visualization of the content

In order to implement a light, wireless and compact prototype, we have selected the OQO handheld computer as the visualization interface. Due to its high performance, this device allows not only the streaming of the personalized multimedia content to visitors, but also the tracking of their position.

The OQO is a compact unit measuring 4.9" x 3.4" x .9", based on a 1GHz Transmeta Crusoe processor that weights approximately 400 g (Figure 9). We found the unit to have a battery life nearly one hour and 45 minutes while requesting web information via wireless LAN, and to be relatively resistant to rough treatment.

Preferences, previous knowledge, interests and movements of the user within the exhibition room have been modelled to achieve a better match between the content and the current context of the user. As content is personalized, visitors must first fill out some simple questionnaires in order to collect some personal details, such as their interests and preferred language.

As mentioned before, when the visitor gets close to a scale model, the visualization application receives the identification of the scale model in front of the visitor. Based on the profile of the visitor, the application makes a query to the server to decide which piece of information is the most appropriate.

Once the piece of information to be shown is determined, the visualization device makes a query to the content database, so that it can create the web site "on-the-fly" and send it to the handheld device. All the multimedia files are stored in the server. Therefore, the only thing that must be installed in the handheld device is the visualization application. This provides the system with high flexibility in order to update contents.

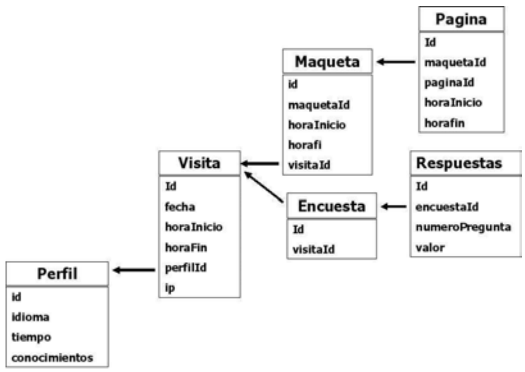


Figure 11: *Schema of the database for the evaluation of the behaviour of the visitor.*

4.6. Evaluation of the behaviour of the visitor

As an additional tool to support the usability study of the AMICo project, the implemented client-server platform stores all the choices made by the user. In order to manage all the information, a third database has been implemented (Figure 11), independent from the information database and the database with the added-value contents generated with the Authoring Tool.

The database stores the profile of the visitor; general data about the visit (data and timetable); the scale models that have been visualized and the visualization time for each content; and the answers to the questionnaires that are filled after the visit.

5. Conclusions and further work

A multimedia information system has been implemented in order to allow users to access specific information related to their immediate surroundings using mobile computing devices. The system has been implemented to provide information about the exhibition of architectural scale models of the famous Spanish architect Rafael Moneo in the Kubo exhibition centre in San Sebatian, Spain. The system allowed the user visualizing images, listening to audio and reading textual information while standing in front of the scale models.

The implemented infrastructure consists of five Access Points (APs) distributed properly so that the whole exhibition room is covered. Each AP communicates with the user's mobile device using Wi-Fi networks. The proposed prototype includes different components, such as a tracking system based on wireless LAN, real time multimedia content streaming via the wireless LAN and personalization of the content due to the definition of different profiles for the users.

Several problems were found regarding the implementation of the tracking system based on the Wi-Fi network, mainly due to the layout of the exhibition room. One problem of the Wi-Fi based location system was the amount of manual calibration effort needed to build the radio map during the offline training phase. It was necessary to compile a fairly dense radio map to compromise a large amount of measurements at different sampled points to attain reasonable position accuracy.

Another issue was that the central exhibition room of the Kubo exhibition centre is too open, so that signal levels were too similar in order to calibrate the system properly. Therefore, the provision of information was limited to three scale models, physically enough far away from each other in the exhibition room to avoid interference, namely the future Church of Riberas of Loyola, the Congress and Exhibition Centre Kursaal, and the Library of the Deusto University.

Further difficulties were found concerning the streaming of the contents in real time. The roaming among the five APs was developed on the basis of existing algorithms to manage the connectivity using the Windows XP driver of the OQOs. In the algorithm, there is no disconnection from one AP even though there could be another AP in the nearby with better coverage of the area. This caused several problems in the efficiency of the streaming of contents in real time, as the signal was lost in some areas of the exhibition room.

Regarding the content, the usability study showed that traditional approaches that have been applied to the design of contents are no longer feasible for current interactive multimedia applications. Therefore, several aspects were highlighted by the users, such as the layout and design of the templates, or the quantity of information provided. However, although different aspects should be improved, users have perceived an added-value in the use of advance mobile guides within the exhibition centre and they would be even willing to rent the guide during their visit for a couple of euros.

Finally, we can conclude that the optimal multimedia guide should support strong personalization of all the information provided in a museum in order to ensure that each visitor is allowed to accommodate and interpret the visit according to his/her own pace and interests. At the same time, a museum guide should foster learning and self-development so as to create a richer and more meaningful experience.

6. Acknowledgement

This paper is part of the work done within the project "AMICo- Asistentes móviles interactivos para la valorización de contenidos digitales", financed by GAIA, Euskalcom and usyscom within the INTEK2004-2005 program of the Department of Industry of the Basque Government. The authors want to also thank the Faculty of Humanities of Deusto University.

References

[AGH*02] AOKI P. M., GRINTER R. E., HURST A., SZYMANSKI M. H., THORNTON J. D., WOODRUFF A.: Sotto voce: Exploring the interplay of conversation and mobile audio spaces. In *Proc. ACM SIGCHI Conf. on Human Factors in Computing Systems* (2002), pp. 431–438.

[CP04] CIAVARELLA C., PATERNO F.: The design of a handheld, location-aware guide for indoor environments. *Personal Ubiquitous Computing 8* (2004), 82–91.

[JEMC05] JAEN J., ESTEVE J., MOCHOLI J., CANOS J.: Momo: enabling hybrid museums. *IEE PRoc. Software 152(5)* (2005), 245–253.

[RS02] R. S., SPASOJEVIC M.: *The Electronic Guidebook: using portable devices and a wireless web-based network to extend the museum experience*. Report HPL-2002-76, 2002.

[RTA05] RAPTIS D., TSELIOS N., AVOURIS N.: Context-based design of mobile applications for museums: A survey of existing practices. In *Proc. of MobileHCI'05* (2005).

[Wil04] WILSON G.: Multimedia tour programme at tate museum. In *Proceedings of the Museums and the Web Conference* (2004).

[ZAS03] ZANCANARO M., ALFARO I., STOCK O.: Using cinematic techniques in a multimedia museum guide. In *Proceedings of the Museums and the Web 2003* (2003).

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Educative Visuals – Digital Delivery of Architectural Information for (potential) Heritage Buildings

Verdy Kwee, Antony Radford and Dean Bruton

The University of Adelaide, Australia

Abstract

The paper proposes models that address current issues and considerations at several key levels relating to treatment of architectural information, its presentation and delivery methods specific to architecture education requirements. It investigates fundamental digital communication strategies for the understanding of architectural work of heritage or potential heritage values, highlighting how digital simulations in particular could complement other media like texts, drawings and photographs to facilitate an understanding of design. It proposes dynamic visual layering system of information and information types relating to site, construction, materials, textures, design philosophy, etc, while also taking into account feedback from the intended audience.

The architectural work featured as an example is of high potential heritage value - an area of special interest in the context of a country with a relatively short architectural history as Australia. The information depicted in the model has a role to supplement a site visit or to communicate independently to the much larger audience who are unable to visit the site.

Although the paper does not insist on definite or final prescriptive techniques for the delivery of architectural information of heritage or potential heritage values, it suggests a possibility of standardisation in this area with features and considerations that need to be firstly addressed.

Categories/Subject Descriptors: E.2 [Data Storage Representations]: Linked Representations; H.5.2 [Information Interfaces and Presentation]: User Interfaces --- User-centred Design; H.5.4 [Hypertext/Hypermedia] Navigation; I.3.6 [Computer Graphics]: Methodologies and Techniques --- Interaction Techniques; I.3.6 [Computer Graphics]: Methodologies and Techniques --- Standard; I.3.7 [Computer Graphics]: Three Dimensional Graphics and Realism --- Animation; I.6.8 [Simulations and Modeling]: Types of Simulations --- Animation; I.6.8 [Simulations and Modeling]: Types of Simulations --- Combined; I.6.8 [Simulations and Modeling]: Types of Simulations --- Visual; J.5 [Computer Applications]: Arts and Humanities --- Architecture; K.3.1 [Computers and Education]: Computer Uses in Education --- Computer-assisted instruction
Additional Keywords: Architectural Information, Media, Architecture Education, Information Representation, Information Visualisation, Computer-assisted Learning, Dynamic Multi-Layered Information

1. Introduction/Background

Information digitisation commands a fast growing interest in the preservation of architectural heritage data. With so much data that have been painstakingly collected, how should they be effectively presented and made available and understood by generations to come? Apart from accessibility and issues concerning message conduits, techniques to achieve coherent deliveries have also become a critically significant issue.

With advances in electronic documentation, graphic presentation and a growing interest in electronic resources, more in-depth information of heritage or potential heritage buildings is likely to be published in digital visual formats and made widely accessible from desktop computers. In order to capitalise on what the present and future technologies offer, there is a need for case studies that explore in-depth types and effective modes of architectural design information representations as an alternative to, and not merely a digital replica of, current print publications.

The main preoccupations of recent research in the digital representation of buildings have been the evaluation and uses of hardware (PDA, Mobile technology, projectors, etc), virtual reality (VR) interactivities, and photo-realistic renderings [AK05, SWL04, NHAH03, YC03, DPF03, WSBW01]. These projects do provide significant factors to consider in the delivery of information content. However, representations and visualisations of architectural heritage work information for those within the academic circle, and especially for the public at large, require attention which cannot be addressed by the improvement of technology alone. Contrary to McLuhan's assertion, [Mcl64] the medium is not the only massage, although it is indeed one of the integral parts of the equation.

Riding on the current available technologies within and outside the field, we approach the investigation to propose some examples of essential delivery techniques. However, it is also realised that in addressing these techniques of delivery, several fundamental concerns specific to architectural information must first be addressed in this paper.

2. Information Types

The type of information pertaining to a building inevitably dictates the manner it is/can be effectively presented. For the purpose of this paper, information types are streamlined into two categories:

- *Hard information*, which might be derived from the subject's physical form, environment and condition; this type of information is often collected directly from the site and partly from architectural documentations.
- *Soft information* (the invisible), which may entail functional and effective observations relating to the realisation of design intentions and their impacts, as well as the building's hidden narratives and phenomenology. This type of information may be less tangible and less visual in nature.

The multi-faceted property of an architectural design's hard information is also echoed in the layers of soft information. This includes the analysis of and narratives about the subject, accounts from users, designers and individuals involved, etc. With heritage buildings or archaeological sites where information resources are limited, this aspect often relies on informed guesses.

3. Media types

Much has been mentioned on the uneasy relationship between language/texts and architectural descriptions (see examples [Fri00, Fro03, Kou97, KRB05]). Although the outcome of this relationship often renders the subject misunderstood and thus the record of knowledge distorted, these forms of print media are widely accepted due partly to their well-understood limitations, portability and ease of use. Similarly, photographs and drawings have limitations but are still necessary. It is not a question of whether they should be used but how they are best used in a complementary manner/relationship.

The effectiveness in delivering certain information types is also partly based on the choice between the visual and the non-visual (literary) media types. In visualising the invisible, for example, Bermudez, Smith & Striefel attempted just that – visualising soft information that relate to the human senses. [BSS05] The resulting visual representations were at best interesting abstract visual metaphors, but they would have been reversibly indecipherable as human sensory experiences.

Since most of the hard information could be relayed visually while the soft information still seems to be more effectively explained verbally or textually, the challenge lies in articulating the interrelationships between the non-visual elements and their visual counterparts in order to establish a referencing system that increases comprehensibility.

4. Information vehicle

A decision on the information conduit requires the assessment of technological opportunities and limitations. This not only involves the facilitation/technology to present and deliver the types of information but also availability of the technology for the intended audience.

Students and the public at large increasingly turn to the Internet as their first and main source of information (refer to figure 1). The use of the Internet presently requires a compromise between image quality and size. What began as an only text-based vehicle has within two decades expanded to delivering images, sounds, and moving pictures [Lev99]. As technology improves, allowing higher compression qualities and bit-rate transfers, it is expected that the delivery of adequate image quality will become mainstream.

2 In studying this architecture, where have you obtained your information from?

Resources	Response Percent	Response Total
Books	62.61	139
Journal(s)/Magazine(s)	55.41	123
Lectures	27.93	62
The Internet/Computer	72.52	161
Video documentary(s)	13.51	30
Others	10.81	24
	Total Respondents:	222
	(skipped this question):	2

Figure 1: Excerpt from the results of an international online survey. [KRBR06-a]

There are also compatibility issues. Operating system variations and the discrepancies found in browsers impact on the display of certain types and format of information. Currently, one of the ways to resolve these incompatibilities is the use of plug-ins, readers or players that work on multiple platforms.

5. Information Contents

It could easily be established that architectural information has undergone various manners of representations in publications- such as architectural monographs (of particular architects' works, etc.), most, if not all of which are author-driven in contents and structures. To whatever extent that these long-established publications may have been relied upon as sources of information and knowledge, there appears to be an absence of assessment of the effectiveness they assume to deliver. It is often found that in such publications, readers are met with unrealistic expectations of prior understanding of the subject matters and the publications themselves often seemed author-centred rather than ones that genuinely guide readers to achieve maximum possible comprehension.

A survey has recently been conducted to investigate information contents that some intended audiences value and consider important to them (Figure 2: blue line) and how existing publications have helped them understand a particular building (Figure 2: red line). (For details, refer to [KRBR06-b]) The chart shows that for most content factors, audience understanding level falls almost consistently below the degree of importance they assign to each factor.

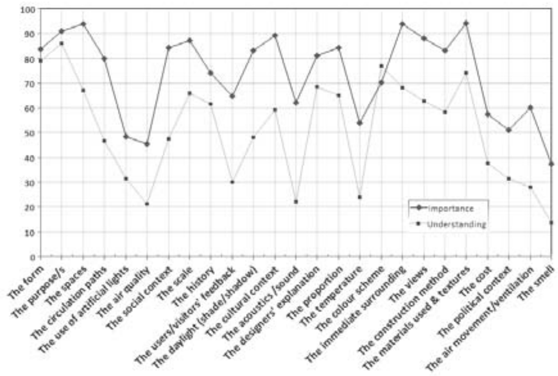


Figure 2: *Audience-surveyed factor importance and understanding scales (based on current publications). [KRBR06-b].*

Although some may argue that this general opinion might have been conditioned by the currently available, author-driven resources, it does provide significant insights into audiences' perceptions of important aspects and of their understanding.

6. Visualisation Standardisation

The documentation of architecture for construction has a long history of commonly-accepted presentation standards. The fundamental and almost universal requirements that have continued over time include the use of two dimensional elevations, plans, sections and details to relay crucial information for the building construction. On the other hand, for descriptions and explanations of already built works in which often a visual third dimension is introduced, there have not been such a broadly understood set of standards. There have been common occurrences of photographs, multi-view perspectives/sectional perspectives, isometric/axonometric views, textured/coloured plans, elevations or sections added to texts. But both the extent of media use and the organisation of materials differ from one published work to another.

Architectural digital presentation is currently in a similar but early stage of non-standardisation. Most examples digitally re-adapt the already non-standard traditional printed mode of delivery (see example: http://www.greatbuildings.com). Where the product is not intended for in-depth public education or detailed assessment this method is tolerable. For instance, architectural illustrations may be regarded and recognised as a form of art, as illustrated in 'Picturing Architecture' [Lus92]. When this happens, wider open interpretations are not only required but desirable.

However, in cases where architectural design knowledge is to be communicated as a record of and educational resource for heritage work, there is value in establishing a yet-to-be-defined fundamental set of visualisation/delivery techniques. Could there exist the possibility of standardisation or a framework to guide visualisation materials meant for this purpose? For this to happen in the digital platform of this era, some modes of delivery need to be devised and/or extended to allow this accessibility and clarity of information. To illustrate and propose a possible direction, examples involving a case study are outlined below.

7. Case Model - Background

One would think that an important aim of preservation of building information of heritage value would be for the benefit and knowledge of the present and future generations. The more pertinent data gathered, the more beneficial the collection would be. It is difficult to fathom, therefore, why preservation works frequently commence only when the information becomes part of the 'Endangered Memory' or is in the brink of the 'Lost Memory' database of an organisation. Early identification and information collection of buildings with heritage potential are essential in order to maintain an acceptable degree of accuracy. It is imperative that documentation commences while untainted and relatively complete information - both hard and soft types - is still available and accessible. This eliminates the need to reconstruct hypothetical scenarios as are often found in cases of heritage or archeological site documentation or reconstruction [HGGS01, Saa01].

In our research, the Arthur and Yvonne Boyd Education Centre, West Cambewarra in the state of New South Wales (NSW), Australia, was selected as a model for data collection and representation mainly for the following reasons:

- It conforms with the local Heritage Council of New South Wales authority's assessment criteria for heritage status [Her01],
- The availability and completeness of information and its absolute significance of the work in the Australian architectural/cultural heritage climate.
- Featured in the Phaidon Atlas of Contemporary Architecture [Pha04], the building is significant to the field as an example of a fine architectural work with sensitivities to the environment.
- It is currently managed by the Bundanon Trust as an education centre for young budding artists to continue the Australian art/cultural lineage. Since it is open to the public, onsite data are readily attainable.
- The availability of people involved to be interviewed. They include the architects, engineers, client representative and builder.
- The relatively remote location of the site necessitates the availability/accessibility of an alternate information source.
- Built on a land donated by the internationally well-known, late Australian painter, Arthur Boyd, it was designed by the only Australian Pritzker-prize winning architect, Glenn Murcutt in association with Wendy Lewin and Reg Lark. The architectural masterpiece, symbolising the efforts by two important figures in the Australian cultural scenes, has won numerous architectural awards, strengthening it as a worthy subject.

8. Visualisation Approaches

Due to the scale of a multi-faceted subject, exemplified in The Arthur and Yvonne Boyd Education Centre, architectural information representations are unusual in that they call for both overall and detailed observations to be clearly included and understood. This frequently translates to a time-demanding process of recording and digital modelling of the subject and more importantly, as outlined later in this paper, unique methods of presenting them faithfully.

Apart from still images, digital presentation vehicles currently in place include predefined walkthrough animations and interactive virtual reality models that can be controlled by the viewer. In architecture, these have mainly been used as a part of the design process for a particular building or to market proposed projects. In academic circles, they are extensively used as illustrations to verbal presentations. As stand-alone education materials, however, they are grossly insufficient.

The following sections highlight two examples of delivery techniques that expand upon the current ones by the introduction of a dynamic information layering system, encompassing various information and types as well as media. Similar concepts could be applied in other visualisation techniques for education purposes.

8.1 Dynamic multi layering approach: Walkthrough / Flythrough Environments

Walkthroughs or flythroughs are useful in providing a quick overall general impression of the site through visual continuity. Typically these are high-quality pre-rendered movies, but they can also involve real-time interactive experiences such as a VR environment [ZdMvG01] or the use of game engines. All these established vehicles have been effective in understanding formal, spatial properties and characteristics. How could we push the boundaries to fashion them into a more effective educational resources?

In the following example, pre-rendered walkthroughs/movies are selected due to:

- the need for focussed information dissemination to facilitate an organised mode of understanding the subject matter, and
- the manipulability and usability of pre-rendered movies in multiple computer platforms, hence facilitating wide accessibility.

Although not as significant as the above factors, another consideration is the quality of images. Even though pre-rendered images may be costly to produce, they take into account global illumination calculations which are still impossible to replicate in real-time digital environments at the time this paper is written. The high quality pre-rendered images therefore, present closer-to-factual data. For more comparative studies, please refer to [SSA01].

A walkthrough or flythrough should ideally attempt to clearly narrate both soft and hard information. One may argue that the narrative/information of a walkthrough will be influenced by the path that the 'camera' is set up to take. As Patrick Keller notes, the narration of an architectural animation sequence is determined by the sequence of visual stills, and not the other way around. "If one had put the pictures together in different order, or had shot different pictures, some other equally plausible fiction might have been the result" [Rat02]. The 'fiction' here can be regarded as narratives about the architecture. Any resulting walkthroughs/paths should be directed in a way that will impart crucial information, although this information will inevitably be selected and incomplete.

The main concern is whether the pre-rendered moving images might miss some important aspects of the building. But this is true with interactive VR environments as well. Most user-controlled VR environments situations lead users to an exercise that emphasizes searching for rather than understanding information; this is aptly is illustrated by IRVE [BNC*] - a project that investigates general ways of representing educational information in such an environment discounting discipline-specific requirements.

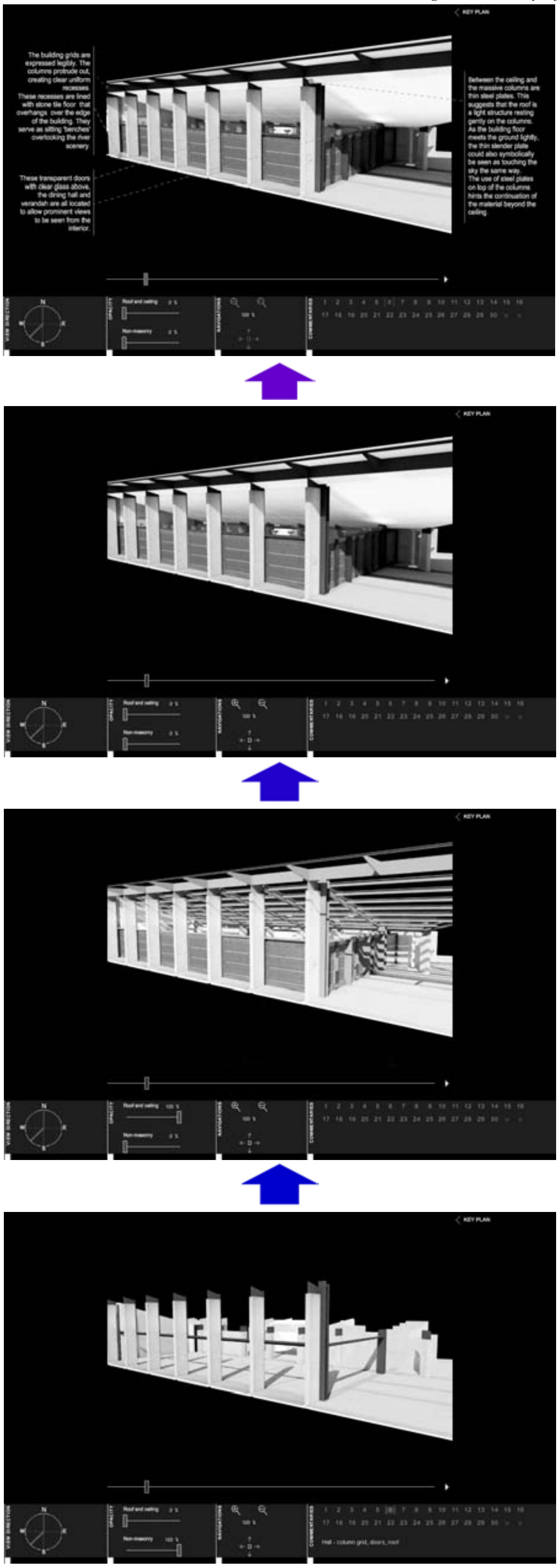


Figure 3: *Screenshots: layering of multiple animations and information.*

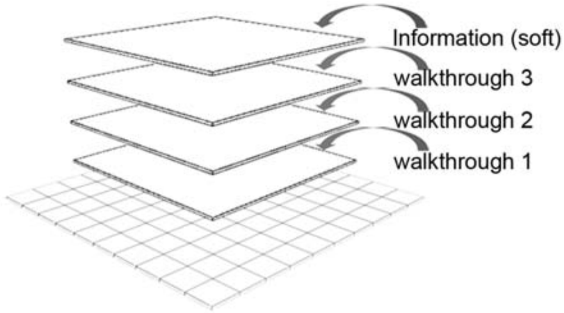


Figure 4: *Stacked information in a walkthrough set with adjustable alpha channels.*

To reduce the possibility of missing crucial information, it is often necessary that several sets of media are available to users, especially in larger scale buildings.

The nature of these moving pictures allows only a limited amount of information to be displayed at a time. Since they bear no true time-space constraint, they may be momentarily paused and relevant, often soft information about that particular scene can be displayed and read. Contextualizing more than one type of media within this also becomes possible. In order to take advantage of this, a separate information layer addresses the limitations of a movie length that might otherwise prevent the delivery of important in-depth narratives. The close interplay between texts and visuals in a visually-continuous environment is significant to contribute to a deeper understanding of the subject and thus helps reduce the level of misinterpretations.

8.2 Dynamic multi layering approach: Referencing to architectural documentation

Traditional sections, plans, and elevations exist in most architectural heritage documents. Possessing symbols and texts that work together well, this information is significant for an in-depth insight into the building composition. They serve as effective platforms to link textual media with graphics symbols to explain a building's construction materials, their locations/arrangements, sizes, etc. However, it does take trained individuals to translate them spatially. To those outside the field, the two-dimensional symbols and texts are hard to understand. Connecting them to rendered models which possess qualities of more decipherable visuals radically increases their legibility. The addition of a third dimension to the traditional two-dimensional drawings assists also in perceiving depth of structural components and spaces.

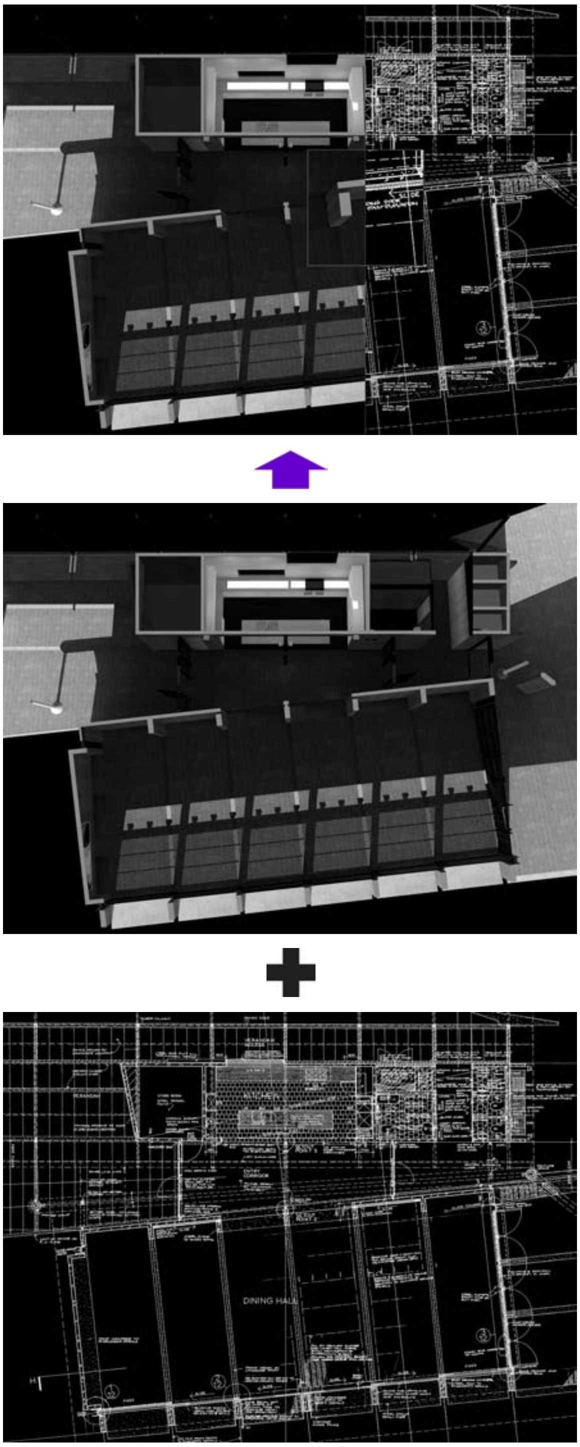


Figure 5: *Interplay of 2D and 3D information to arrive at an interactive, visual plan of hall with enlarged portion (in red-bounding box) for increased resolution/readability.*

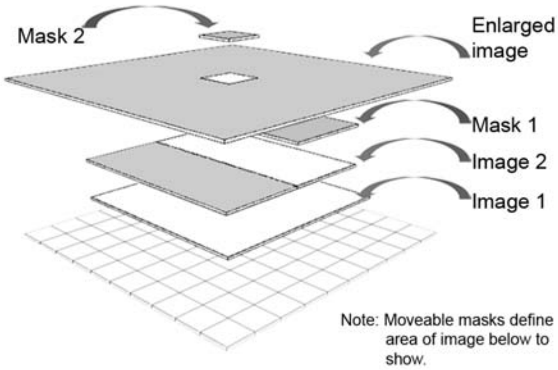


Figure 6: *Stacked still images of close reference.*

Considerations need to be taken in reference to the legibility of the document during the digitisation process and the screen-size in which they are finally presented.

9. Conclusions

The power of architectural publications is often taken for granted and easily underestimated or overlooked in the quest for achieving an educational impact. This is exhibited in a compilation of articles in 'Educating Architects' where these publications' presence and contributions have clearly not been acknowledged.[PT95] Digital delivery could alter this perception and augment the role of such publications to be an inseparable entity of architecture education.

More than the traditional publications, trends indicate that the public is leaning towards readily accessible electronic information when architectural information is concerned. As the internet-based survey results suggest, digital media are approaching and even surpassing printed media as a significant source of information. This is obviously a concern when currently the reliability and substance found in print media still exceed those of the widely accessible digital format.

The ideas presented in the model here are intended to promote interest in the capabilities of technologies in delivering architectural information. There are indeed myriads of features and information to be exploited in order to facilitate better understanding of architectural designs and especially those of high heritage values by extracting and expanding techniques from within the architectural field as well as from outside the industry. The examples proposed in this paper thus, are not meant to be the final and only solutions.

As hinted at by Paraizo [Par04], the possibility of visualisation techniques are limited only by the imagination and suitable use of 'hyperdocumentation'; these techniques are of course, discipline-specific. In his digital visualisation of urban structure of Rio de Janeiro, Ripper Kós [1999] has adopted this hyperdocument approach of combining a selection of still pictures to deliver the information of the city. Given the nature and details of architectural information, however, the execution and choice of media would unquestionably be different if it were to narrate the information of a building that exists within that urban fabric.

Some have observed that there has been no evaluation of the majority of proposed visualisation systems [RB05]. Others may also counter-argue that while there is an absent of valid, reliable testing instruments to start with, for approaches like those shown here to be a part of the future visualisation mainstream will realistically take more than a one-time validity testing to be successfully implemented. At present, technologies are still incessantly reconfigured and consequently, visualisation systems are likely to morph and improve. Thus, a 'validity' testing would prove age/time specific, while the data falling into the danger of being subscribed beyond their likely short expiry date. Due to this dependency on other factors, visualisation systems can/should not be deemed final or finalised by evaluated 'confirmation' at this point in time but considered as steps and influences towards better forms and facilitations.

That being said, the emergence of more effective common delivery modes for architectural information for the academics and the public at large is optimistically foreseeable as facilitating technologies further develop and perhaps plateau. Further studies of techniques of architectural information delivery and much experimentation need to work in tandem with available technologies both to take advantage of the opportunities and set the direction for a future seamless technological integration. Further convergence of those technologies overseen and capitalised by the content creators may indeed help better facilitate the understanding of architectural heritage, potential heritage and archaeological excavation projects.

Acknowledgments

We thank:

- The Bundanon Trust for access to The Arthur and Yvonne Boyd Centre to facilitate extensive onsite study, recording and measurements.
- The architects Glenn Murcutt, Wendy Lewin and Reg Lark as well as Mitchell Library for access to the architects' drawings and other priceless information.
- James Taylor and Associates for access to the engineers' drawings.
- Sue Barnsley, Peter Bacon, Jonko Berg for information on landscape, onsite systems and development processes.
- All individuals around the globe who have participated in our online survey and/or helped in disseminating the announcement.
- Friends/colleagues who have provided invaluable advice and assistance.

References

[AK05] AFFLECK J., KVAN T.: Reinterpreting Virtual Heritage *CAADRIA 2005 Proceedings of the 10th International Conference on Computer Aided Architectural Design Research in Asia*, New Delhi, India, vol. 1, (April 2005) pp. 169-178. http://cumincad.scix.net/data/works/att/caadria2005_a_2b_a.content.pdf.

[BSS05] BERMUDEZ J., SMITH A. C., STRIEFEL S.,: Visualizing the Invisible: Digital Studies on Representing Non-Visual Architectural Experiences, *CAADRIA 2005 Proceedings of the 10th International Conference on Computer Aided Architectural Design Research in Asia* New Delhi, India, vol. 2, (April 2005), pp. 510-515. http://cumincad.scix.net/data/works/att/caadria2005_b_6c_e.content.pdf.

[BNC*03] BOWMAN D. A., NORTH C., CHEN J., POLYS N. F., PYLA P. S., YILMAZ U.: Information-rich virtual environments: theory, tools, and research agenda, *2003 VRST* (October 2003), 81-90. http://portal.acm.org/citation.cfm?doid=1008653.1008669.

[DPF03] DUMONT R., PELLACINI F., FERWERDA J. A., Perceptually-driven decision theory for interactive realistic rendering. *ACM Transactions on Graphics*, vol 22/2, (2003), pp 152-181. http://portal.acm.org/citation.cfm?id=636888&coll=GUIDE&dl=GUIDE&CFID=112014&CFTOKEN=69813742.

[Fri00] FRIEDMAN A.: Language and Movement in CAD Application. *CAADRIA 2000 Proceedings of the Fifth Conference on Computer Aided Architectural Design Research in Asia*, Singapore, 18-19 (May 2000), pp. 423-432. http://cumincad.scix.net/data/works/att/3cd1.content.pdf.

[Fro03] FROSH P., *The Image Factory,* (2003), pg 19-23.

[FST92] FUNKHOUSER T. A., SÉQUIN C. H., TELLER S. J.: Management of large amounts of data in interactive building walkthroughs. *Proceedings of the 1992 symposium on Interactive 3D graphics*, (June 1992). http://portal.acm.org/citation.cfm?id=147158&dl=ACM&coll=portal&CFID=15151515&CFTOKEN=6184618.

[Her05] HERITAGE OFFICE, *Assessing Heritage Significance.* New South Wales, 2001. http://www.heritage.nsw.gov.au/03_index.htm.

[HGGS01] HYNST S., GERVAUTZ M., GRABNER M., SCHINDLER K.: A work-flow and data model for reconstruction, management, and visualization of archaeological sites. *Proceedings of the 2001 conference on Virtual reality, archaeology, and cultural heritage*, (2001), pp.43-52. http://portal.acm.org/citation.cfm?id=585000&coll=GUIDE&dl=GUIDE&CFID=112014&CFTOKEN=69813742.

[Kou97] KOUTAMANIS A.: Digital Architectural Visualization. *Challenges of the Future: 15th Ecaade Conference Proceedings*, Vienna, Austria, (September 1997). pp.17-20. http://cumincad.scix.net/data/works/att/cb26.content.pdf.

[KRB05] KWEE V., RADFORD A., BRUTON D. Hybrid Digital Media Architectural Visualisation Delivery -Murcutt, Lewin & Lark's The Arthur and Yvonne Boyd Education Centre on Digital Flatland. *International Journal of Architectural Computing*. Vol 3 (4), (Dec 2005) pp 487-502.

[KRBR06-a] KWEE V., RADFORD A., BRUTON D. ROBERTS I.: *Architecture | Media| Representations: Survey Data 2006*. (2006) http://cumincad.scix.net/cgi-bin/works/Show?8d88.

[KRBR06-b] KWEE V., RADFORD A., BRUTON D. *Architecture | Media| Representations: A Reassessment.* Paper to be presented at ANZAScA 2006, Adelaide, Australia, (22-25 Nov 2006).

[Lev99] LEVINSON P.: *Digital McLuhan: A Guide to the Information Millennium*: New York: Routledge (1999). p.5.

[Lus92] LUSCOMBE D. O., PEDEN A.: *Picturing Architecture: Graphic Presentation Techniques in Australian Architectural Practice*. Tortola, British Virgin Islands ; Roseville East, N.S.W: Craftsman House, (1992).

[Mcl64] MCLUHAN M.: *Understanding Media : The Extensions of Man*. London: Routledge & Kegan Paul, 1964.

[NHAH03] NIEDERAUER C., HOUSTON M., AGRAWALA M., HUMPHREYS G.: Non-invasive interactive visualization of dynamic architectural environments, *Proceedings of the 2003 symposium on Interactive 3D graphics*, (2003), pp55-58. http://portal.acm.org/citation.cfm?id=882331&coll=GUIDE&dl=GUIDE&CFID=112014&CFTOKEN=69813742.

[Par04] PARAIZO R. C.: 2004. A representação do patrimônio urbano em hiperdocumentos: um estudo sobre o palácio monroe [The Representation of Urban Heritage in Hyperdocuments*:* A Study on the Monroe Palace], *SIGraDi 2004 - [Proceedings of the 8th Iberoamerican Congress of Digital Graphics]* Porte Alegre - Brasil (November 2004), 10-12. http://cumincad.scix.net/data/works/att/sigradi2004_200.content.pdf.

[Pha04] *The Phaidon Atlas of Contemporary World Architecture*, Phaidon Press, London; New York, (2004), p.39.

[PT95] PEARCE M., TOY M.: *Educating Architects*: London: Academy Editions, (1995).

[Rat02] RATTENBURY K.: *This Is Not Architecture : Media Constructions*. London ; New York: Routledge, (2002). p.40.

[RB05] RIPPER KÓS, J., BARKI, J.: Between text and hypertext: a case in architectural theory, *SIGraDi 2005 - [Proceedings of the 9th Iberoamerican Congress of Digital Graphics]* Lima - Peru 21-24 (November 2005), vol. 2, pp. 788-792.

[Rip03] RIPPER KÓS, J., (2003). *Rio-H [online]* http://www.kos.med.br/thesis.

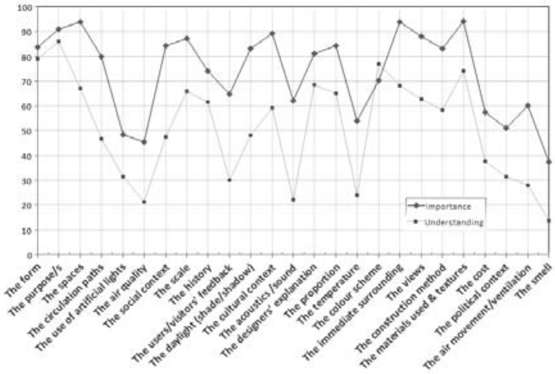
[Saa01] SASS L.: Reconstructing Palladio's Villas: A Computational Analysis of Palladio's Villa Design and Construction Process. In Reinventing the Discourse - How Digital Tools Help Bridge and Transform Research, Education and Practice in Architecture. *Proceedings of the Twenty First Annual Conference of the Association for Computer-Aided Design in Architecture*. Buffalo (New York) 11-14 (October 2001), pp. 212-226. http://cumincad.scix.net/data/works/att/5166.content.pdf.

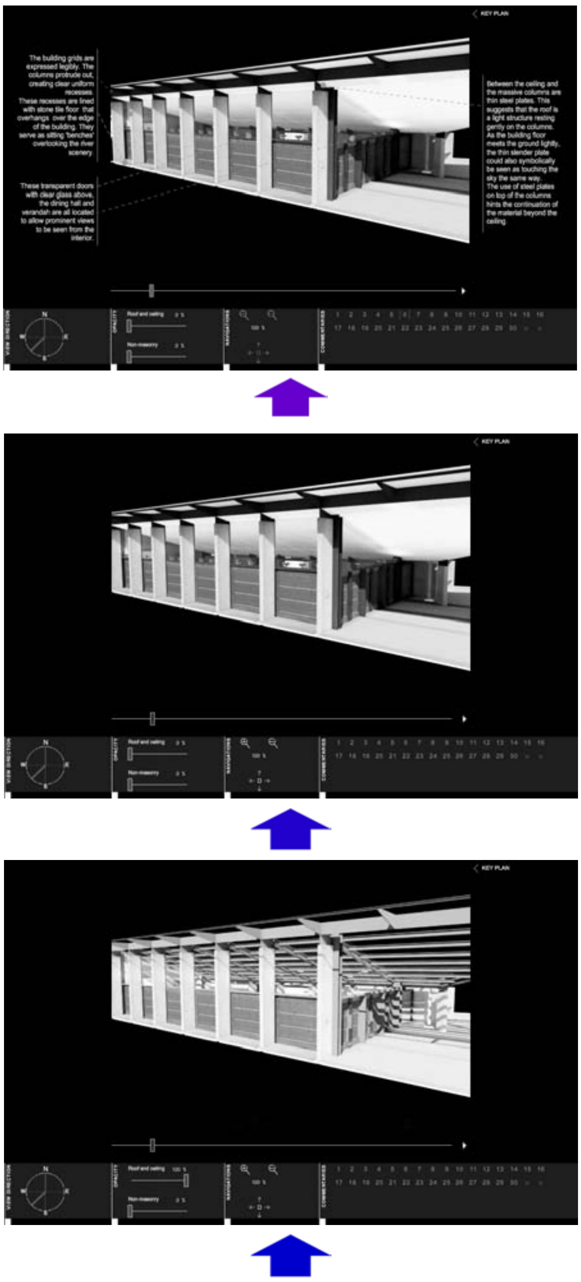
[SSA01] STAPPERS P. J., SAAKES D., ADRIAANSE J.: On the Narrative Structure of Virtual Reality Walkthroughs. An Analysis and Proposed Design. *In Proceedings of the Ninth International Conference on Computer aided Architectural Design Futures,* Eindhoven, 8-11 (July 2001), pp. 125-138. http://cumincad.scix.net/data/works/att/3875.content.pdf.

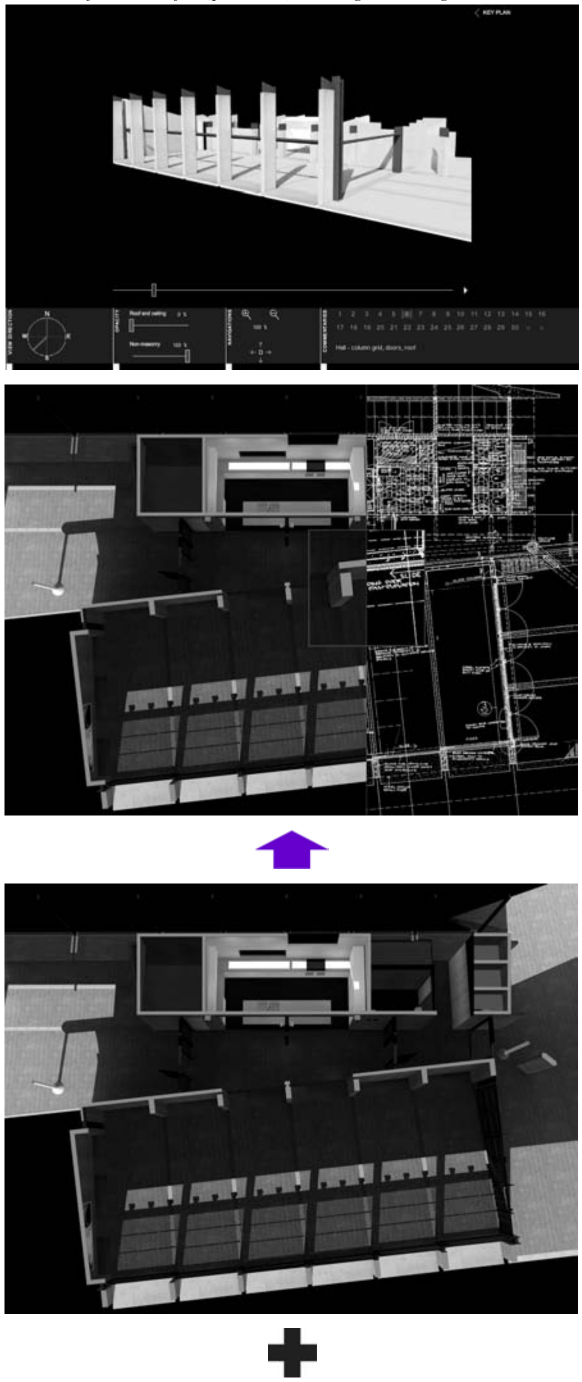
[SWL04] SIU A. M. K., WAN A. S. K., LAU R. W. H., Modeling and Rendering of Walkthrough Environments with Panoramic Images. *Proceedings of the ACM symposium on Virtual reality software and technology* (2004), p 114-121. http://portal.acm.org/citation.cfm?id=1077557&coll=GUIDE&dl=GUIDE&CFID=112014&CFTOKEN=69813742.

[WSBW01] WALD I., SLUSALLEK P.: BENTHIN C., WAGNER M.: *Interactive Rendering with Coherent Ray Tracing.* EUROGRAPHICS (2001), vol 20-3. http://www.eg.org/EG/CGF/volume20/issue3/508.pdf.

[YC03] YANG L., CRAWFIS R.: Panoramic Walkthrough System with Occlusion Culling, Proceedings of the workshop on Virtual environments (2003), pp 169-177. http://portal.acm.org/citation.cfm?id=769973&coll=GUIDE&dl=GUIDE&CFID=112014&CFTOKEN=69813742.


Importance
Understanding
The form
The purpose/s
The spaces
The circulation paths
The use of artificial lights
The air quality
The social context
The scale
The history
The users/visitors' feedback
The daylight (shade/shadow)
The cultural context
The acoustics/sound
The designers' explanation
The proportion
The temperature
The colour scheme
The immediate surrounding
The views
The construction method
The materials used & textures
The cost
The political context
The air movement/ventilation
The smell


KEY PLAN
The building grids are expressed legibly. The columns protrude out, creating clear uniform recesses.
These recesses are lined with stone tile floor that overhangs over the edge of the building. They serve as sitting 'benches' overlooking the river scenery.
These transparent doors with clear glass above, the dining hall and verandah are all located to allow prominent views to be seen from the interior.
Between the ceiling and the massive columns are thin steel plates. This suggests that the roof is a light structure resting gently on the columns. As the building floor meets the ground lightly, the thin slender plate could also symbolically be seen as touching the sky the same way. The use of steel plates on top of the columns hints the continuation of the material beyond the ceiling.
KEY PLAN
KEY PLAN


KEY PLAN

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Geographical Presentation of Virtual Museum Exhibitions

M. Stawniak and K. Walczak

Department of Information Technology, The Poznan University of Economics
Mansfelda 4, 60-854 Poznan, Poland , +48 61 848 05 49
[stawniak, walczak]@kti.ae.poznan.pl

Abstract

In this paper we present a system, called GeoARCO, which enables presentation of virtual museum exhibitions in a geographical context. The system is partially based on the results of the European project ARCO — Augmented Representation of Cultural Objects, which has developed technology for museums enabling them to create and manage virtual museum exhibitions for use in interactive kiosk displays and on the Web. GeoARCO uses the Google Earth platform to enable presentation of digital artefacts as well as complete cultural heritage exhibitions on top of the 3D globe model. Users can browse and search available exhibitions, display current location of objects as well as historical data about the objects, such as the place where the objects were made or discovered. A user can also display detailed 3D models of artefacts, reconstructed sites or entire virtual exhibitions. The system cooperates with multiple ARCO databases run by different museums.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism; H.2.8 [Database Management]: Database Applications - Spatial databases and GIS; K.3.1 [Computers and Education]: Computer Uses in Education

1. Introduction

Recently we can observe how virtual reality and augmented reality technologies are maturing from an early experimentation phase to a practical deployment phase, where 3D VR and AR applications become more and more popular in various domains. This is partially a result of enormous progress in hardware performance — especially graphics cards, increasing availability of automatic 3D modelling tools [SKZ05], and development of standards such as VRML/X3D [Web] and MPEG-4 [Koe02] enabling platform-independent representation of interactive 3D contents. Furthermore, wide availability of 3D computer games and movies based on 3D computer graphics results in increasing familiarity of users with 3D graphics and — at the same time — is raising user expectations. Young generations of people familiar with interactive 3D games and movies based on 3D graphics look for similar experiences in other domains.

One of prominent application areas of virtual and augmented reality technologies is cultural heritage. Museums around the world hold countless artefacts that they cannot exhibit to the public due to limited space, items' fragility, or the prohibitive cost of creating and managing appropriate displays. Travelling exhibitions or those that draw on multiple collections are even more problematic because of the expenses associated with transporting and insuring priceless objects, along with museums' reluctance to part with certain treasures. Yet such exhibitions typically have the widest appeal and generate the most revenue.

Virtual reality and augmented reality technologies offer museums and other cultural heritage institutions a convenient alternative method of presenting cultural heritage resources — both existing and reconstructed. The resources can be presented in the form of highly educational and entertaining interactive scenarios.

The use of VR and AR technologies is particularly beneficial, when the digitized collections of objects are accessible remotely over the Internet. Such virtual museum exhibitions on the web enable different audiences, including the disabled and students of all ages, as well as the general public to remotely access and interact with vast numbers of objects scattered among various localities in an engaging and informative way. Such exhibitions, however, suffer from the lack of one important element — the geographical context.

In this paper, we present a system, called GeoARCO, that enables presentation of virtual museum exhibitions in the geographical context. The system is partially based on the results of the European project *ARCO — Augmented Representation of Cultural Objects* [ARC], which has developed technology for museums to enable them to create and manage virtual museum exhibitions for use in local interactive kiosk displays and on the Web. GeoARCO uses the Google Earth platform to enable localized presentation of digitised artefacts as well as complete cultural heritage exhibitions on top of the 3D globe model. Users can browse and search available exhibitions, display current location of objects as well as historical data about the objects, such as the place where the objects were made or discovered. A user can also display detailed 3D models of artefacts, reconstructed sites or entire virtual exhibitions. The system can cooperate with multiple ARCO databases run by different museums.

The rest of this paper is organized as follows. In Section 2 related works are shortly described and rationale behind our research is presented. Section 3 provides an overview of the ARCO system describing its goals and capabilities. In Section 4, the GeoARCO platform is presented in details. Section 5 provides examples of interactive GeoARCO presentations. Finally, Section 6 concludes the paper and indicates future works.

2. State of the art and rationale

In the field of cultural heritage, digital mapping and geographical data management through Geographic Information Systems (GIS) is long recognised as an important element for conservation, protection, analysis, communication and learning purposes. Recently, numerous free or commercial web-enabled GIS software packages have been developed, which have led to many successful implementations of Internet GIS-driven services and maps, such as *Electronic Cultural Atlas Initiative* [Eleb] [Elea]. With increasing graphical capabilities of personal computer systems and widening access to broadband Internet, 3D GIS visualisation applications become more and more popular, with the *Skyline* [Sky], *Google Earth* [Goob], and *NASA World Wind* [Nat] being prominent examples. Google Earth together with its 2D version — *Google Local* [Gooc] — made a breakthrough in the availability of GIS visualisation services to the broad public.

There has been a lot of effort in employing Google Earth application in the cultural heritage domain. Individuals and organisations across the world are developing collections of historical places, monuments and objects that can be visualised within Google Earth. In most cases such information has a form of static datasets, which are created once and contain a number of points on the map with descriptions, pictures and sometimes also links to corresponding web sites. The examples here are cultural heritage sites in *New Zealand and Australia* [New], *National Geographic Africa Magaflyover* [Wil] and *National Geographic ZipUSA* features [BB].

There have been also several attempts to integrate Google Earth with databases of cultural resources. In this approach the information from the database is transformed into a special XML file that can be downloaded, read by Google Earth and visualised. The Google Earth version of *UNESCO World Heritage Map* [Uni] is a good example.

Since Google Earth is able to display 3D models on top of the Earth surface, the application can be also used for visualisation of 3D reconstructions of heritage sites, especially those with archaeological or architectural focus. *Google Earth Blog* [Tay] and *Google 3D Warehouse* [Gooa] feature many of such objects grouped into several categories.

The examples presented above, although stimulating, are in fact very limited. Most of them use simple 2D web contents to present the heritage resources, which in practice restricts the presentation to textual descriptions and photographs. The systems that use 3D representation of objects or sites, try to visualize them directly embedded within the 3D GIS system (e.g., in the Google Earth). We argue that such approach is not appropriate or at least not optimal because of several problems:

1. The quality of 3D models that can be currently displayed directly within the GIS platforms is low. Introduction of the last version of Google Earth (4.0) has improved the presentation capabilities [Col], but still this is not enough for faithful representation of digitized museum artefacts or architectural designs.
2. Direct embedding of 3D objects into the GIS platform requires conversion of their digital representations into a system specific format (e.g., KML or COLLADA in case of Google Earth platform). This format cannot be used in other presentations of the contents — either local or on the web.
3. In case of digital representation of artefacts, the paramount problem is related to scale. The artefacts must be represented in a different scale than the geographical information. When looking at objects a user looses the geographical context, when looking at the geographical information, the user cannot see the objects.
4. Another problem is caused by distribution of objects. When the system tries to directly position objects in their origin (or some other related location), the objects are distributed at a large area making it difficult for users to locate the objects, navigate between them, compare objects, etc. As a result, most advantages of 3D virtual museum exhibitions are lost.
5. Presentation capabilities of popular GIS platforms (such as Google Earth) are usually limited to displaying the outside surface of objects, not enabling to visualize interiors. This may be sufficient for simple representation of shapes of buildings, but is not sufficient to represent architectural interiors, e.g. a reconstruction of a museum room. Also, the navigation capabilities are limited, e.g. not enabling to view the object from below.

6. The last — but important — problem is related to interactivity. 3D presentation of museum artefacts enables building interactive scenarios, which have the potential to significantly enrich user experience while learning about the objects or places. This is not possible (at least currently) when the objects are directly embedded within the GIS platform.

To avoid the above mentioned problems and limitations, we have decided to use a different approach. We have integrated ARCO — a system able to build high-quality, interactive, standards-compliant virtual museum exhibitions — with a popular GIS platform (Google Earth), taking the best of the two worlds. ARCO enables interesting and engaging presentation of the digitized contents, while the GIS platform displays, in a user-friendly way, all geographical information related to the digitised collections. These two subsystems can communicate with each other while a user is interacting with the presentation.

3. The ARCO system

3.1. Overview of ARCO

The recently completed European Union funded research project *ARCO — Augmented Representation of Cultural Objects* [ARC] [WCW06] has developed technology for museums to create and manage virtual exhibitions for use on the web and in interactive kiosk displays. The ARCO system comprises simple to use authoring tools, which allow exhibition designers to set up virtual exhibitions in just a few minutes. The exhibitions can be then presented both inside museums, e.g. on touch-screen displays installed inside galleries and at the same time on the Internet. The presentation is based on virtual and augmented reality technologies as an extension of standard web presentations and allows visitors to interact with the contents in an intuitive and exciting manner.

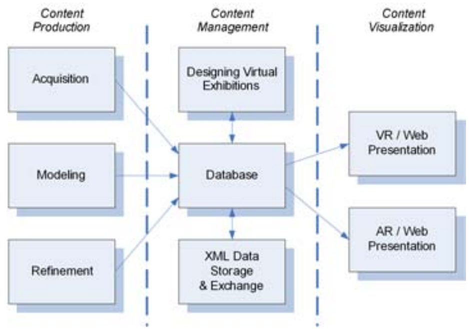


Figure 1: *Overall architecture of the ARCO system*

The overall architecture of the ARCO system is presented in Figure 1. The ARCO system consists of three main architectural components: *content production*, *content management*, and *content visualization*. The content production includes all tools and techniques used to create digital representations of museum artefacts. The digital representations are stored in a database and managed with the *ARCO Content Management Application — ACMA*. Each digitized artefact stored in the database is called a *cultural object*. Each cultural object is represented as a set of *media objects* and associated *metadata*. Examples of media objects are images, movies, 3D VRML/X3D models, QuickTime VR images, descriptions, and sounds. Metadata is structured according to *AMS — ARCO Metadata Schema* [PWM*05].

In addition to managing digitized cultural objects, the ACMA tool enables building interactive virtual exhibitions that present collections of cultural objects retrieved from the database. Using the *ACMA Presentation Manager* a user can build virtual exhibitions by creating exhibition spaces and assigning cultural objects and presentation templates to the spaces. By using advanced presentation templates, interactive presentation scenarios can be built, e.g. implementing learning scenarios associated with the cultural content. Data can also be exported from the ARCO database into an XML data format to move contents between databases and to setup prearranged exhibitions.

The visualization of the virtual exhibitions may be performed by a standard web browser or a web-based augmented reality (AR) application. The AR application integrates a web browser and an AR browser [KB] allowing presentation of either 2D/3D web contents or AR virtual exhibitions. The classical web browser interface allows users to search and browse the database contents by the use of a well-known interface, whereas the VR and AR exhibitions let them examine virtual reconstructions of selected objects in 3D environments.

The virtual exhibitions displayed in the VR and AR interfaces are dynamically generated based on the database contents — the exhibition spaces, the presentation templates and the cultural object models. The use of different templates enables different presentation of the same content. Different forms of exhibitions can be also achieved by the creation of template instances derived from the same template but supplied with different sets of parameter values. The parameters may affect both the visual and behavioural elements of the exhibitions. The virtual exhibitions can be customized for a particular user or created in response to a user query.

ARCO uses X-VRML for creating parameterized templates of virtual exhibitions [WC02] [WC03]. X-VRML is a high-level XML-based language that adds dynamic modelling capabilities to virtual scene description standards such as VRML and X3D. The dynamic modelling technique enables the development of dynamic database-driven VR applications by building parameterized models (templates) of virtual scenes that constitute the application, and dynamic generation of instances of virtual scenes based on the models, data retrieved from a database, current values of model parameters, input provided by a user, and user privileges or preferences. In ARCO, the X-VRML templates define both the visual and the behavioural aspects of virtual exhibitions. The visualization is performed using standards such as VRML and X3D. For the behavioural aspects, high-level XML scripts are used.

The use of presentation templates enables separation of the processes of preparing contents (the cultural objects), designing the presentation form (the templates), and setting up virtual exhibitions, allowing the latter to be easily performed by users without extensive knowledge in computer science and 3D technologies. All the visualization and interaction rules necessary to build exhibitions are encoded in the templates. An exhibition designer can create an interactive exhibition by simply collecting the object models, setting their visualization and interaction parameters and creating an instance of a presentation template. The process of designing an interactive presentation can be performed by the use of a simple application connected to the database.

3.2. Designing virtual exhibitions

The structure of ARCO virtual exhibitions is determined by a hierarchy of exhibition spaces stored in the database. The exhibition spaces are conceptually similar to folders that may contain three types of elements:

- *presentation template instances*,
- *cultural object instances*, and
- *cultural object selection rules*.

A template instance is a template supplied with actual values of some of its formal parameters. A single template can have an arbitrary number of instances in different exhibition spaces, which provides high flexibility in designing virtual exhibitions because different template instances, which imply different visualization and interaction parameters, can be set for every exhibition space. Flexible assignment of parameter values to template instances makes it possible to easily combine search interfaces, customizable browsing interfaces, and fixed exhibitions.

A cultural object instance is a cultural object together with object's presentation parameter values. Again, the same object may have instances in more than one exhibition space. A cultural object selection rule is a search statement that retrieves from the database all objects that meet criteria defined in the rule. The selection is performed when a user accesses the exhibition space and thus enables building virtual exhibitions that are always up-to-date.

An exhibition designer can provide the template parameters and cultural object parameters while setting up a presentation (see Figure 2). When a user enters an exhibition space, all cultural object instances that are assigned to this particular exhibition space or retrieved by a selection rule are displayed in a way defined by the presentation template instance assigned to the same space.

The same exhibition space may be displayed differently in different environments by the use of different presentation templates. To achieve maximum flexibility with respect to different presentation methods, the concept of presentation domains has been introduced. A presentation domain is the context in which the exhibition is intended to be used (e.g., local web environment, remote web, augmented reality scene, or a GIS system). Each presentation template is associated with a list of allowed presentation domains, but each template instance corresponds to a single domain. In an exhibition space, multiple instances of templates for different domains may be created, but at most one instance for each domain. While accessing an exhibition, the browser specifies which domain should be used. Then, the appropriate instance of the template is used to dynamically produce the contents.

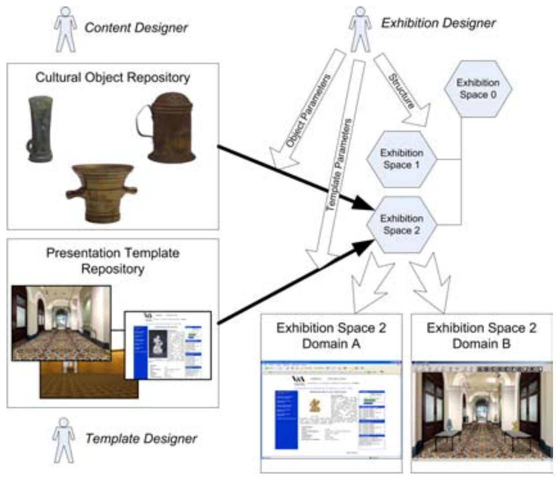


Figure 2: *ARCO — building virtual exhibitions*

In order to speed-up the process of designing exhibitions and to ensure consistency of presentation of objects organized in the exhibition spaces hierarchy, the concept of inheritance of template instances has been introduced. In this approach, if a specific exhibition space does not contain its own template instance in the indicated presentation domain, the instance contained in its parent space is used by default (recursively). This solution enables using a single template instance for the whole tree of spaces in the exhibition hierarchy and thus saving the preparation time and ensuring visual and behavioural consistency of presentation within a virtual exhibition.

4. The GeoARCO system

4.1. Overview and architecture

The GeoARCO system is an extension of ARCO enabling geographical presentation of cultural objects and virtual exhibitions. GeoARCO can be used in the same way as ARCO, but, additionally, the Google Earth client can be used to browse and display the virtual exhibitions in a geographical context.

Google Earth is a free application that allows users to browse aerial photographs of the entire planet. An important feature of this application is that on the top of the aerial photographs it can display other kinds of information, such as image layers, points of interest, lines, texts and even simple 3D models. The elements displayed on the map can contain hyperlinks, which can be opened in an integrated web browser.

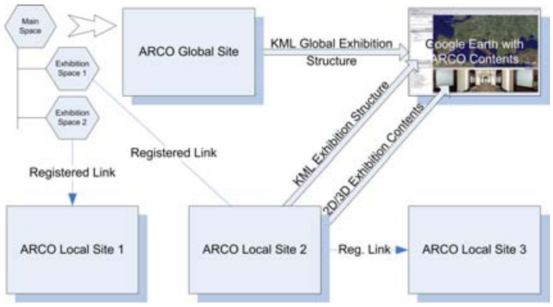


Figure 3: *Architecture of the GeoARCO system*

Figure 3 presents the overall architecture of the GeoARCO system. There are four instances of the ARCO system shown in this example. The instances may be run and maintained by different museums. Exhibition spaces in the GeoARCO system, in addition to the elements described in the previous section, can contain links to exhibitions in other instances of the ARCO system. In such a way, the ARCO instances become interconnected similarly as web servers. The process of linking is performed by an exhibition designer who may include in an exhibition space other spaces available in different museums (i.e., in other instances of ARCO).

The network of ARCO instances is flat in the sense that it does not form a fixed hierarchy. The hierarchy of spaces is dynamically defined by a hierarchy of exhibition spaces in the entry ARCO instance and the connected instances. Depending on the entry point, the same server may be used as a global site or any of the sub-sites in the dynamic hierarchy.

When a user connects to an ARCO site using the Google Earth application, based on the structure of the exhibition spaces, a hierarchy of *Places* is generated and transferred to the application in the KML format [Goo06]. The hierarchy contains the entry space, all subspaces (both local and linked from other ARCO instances) and all objects assigned to the spaces. For each item, the name and basic description are provided. Clicking on an item in the hierarchy, a user can navigate to the item's location on Earth, and display its picture and description together with additional information including links to object's 3D model and object's representations in virtual galleries (either 2D or 3D). The 3D model and virtual galleries can be displayed in the embedded browser (see Section 5).

4.2. Geographical data in ARCO

In order to properly represent geographical data related to cultural objects we have extended the XML schema of metadata used for describing cultural objects (AMS). The additional geographic information includes the coordinates of the object's origin, place where the object is currently located, place where the object has been temporarily moved and other places related to the object.

In addition to the cultural object metadata, coordinates may be also specified in the presentation parameters and assigned to either cultural objects or exhibition spaces. This feature is useful for positioning virtual exhibitions in specific places on the Earth model. Also, in some cases it is useful to position objects around an exhibition regardless their real origin or current location.

The third element that influences geographical presentation of objects and virtual exhibitions is the "belongs-to" relationship. This relationship is specified by adding cultural objects or sub-spaces to exhibition spaces in the ARCO Presentation Manager. The relationship is then represented as connecting lines on the 3D model of Earth (see Section 5).

4.3. Dynamic KML

As indicated in the Section 4.1, in GeoARCO, the Google Earth application is controlled by the use of dynamically generated KML contents. KML is a language based on XML, which allows users to define what information and how is to be presented on top of the Earth model. The GeoARCO KML descriptions are generated based on the structure of exhibition spaces and cultural objects stored in the ARCO databases.

In Figure 4, an example of a KML description of an exhibition space generated by the GeoARCO system is presented. The generated KML file contains a *Folder* node, which represents an exhibition space in the ARCO system. The *Folder* node can contain a number of *Placemark* nodes. A placemark in Google Earth is represented as an icon with a label positioned at some specific coordinates on the globe. Additional description can be attached to a placemark. In GeoARCO, the first *Placemark* represents the exhibition space itself. The remaining *Placemark* nodes represent cultural objects assigned to the exhibition space. There is one *Placemark* corresponding to a cultural object in the presented example.

Sub-spaces are represented by the *NetworkLink* nodes. A network link enables inclusion of any other KML file in the current KML file. This approach has two important advantages. Firstly, it enables to easily reflect the recursive nature

```
<kml xmlns="http://earth.google.com/kml/2.1">
  <Folder>
    <Placemark>
        <name>Art in Western Poland</name>
        <description>...</description>
        <Point>
            <coordinates>16.83688,53.05467,0</coordinates>
        </Point>
    </Placemark>
    <Placemark>
      <name>Woman with butter churn</name>
      <description>...</description>
      <Point>
        <coordinates>16.97409,51.76316,0</coordinates>
      </Point>
    </Placemark>
    <NetworkLink>
      <name>Art in Western Poland</name>
      <visibility>1</visibility>
      <Url>
        <href>http://...</href>
        <refreshMode>onInterval</refreshMode>
        <refreshInterval>60</refreshInterval>
      </Url>
    </NetworkLink>
    ...
  </Folder>
</kml>
```

Figure 4: *Example of a KML specification*

of the exhibition spaces hierarchy in the KML files. Secondly, network links enable automatic refreshing. Therefore, any changes made in the ARCO database can be automatically reflected in the Google Earth application in a short time.

To enable automatic generation of KML contents an appropriate X-VRML template has been developed. The template is assigned to a special presentation domain. The template retrieves objects and sub-spaces assigned to an exhibition space, extracts their position information and displays them in the proper places on the globe. This template may have instances in all exhibition spaces along with 2D or 3D presentation templates assigned to other domains such as WEB_LOCAL, WEB_REMOTE, AR, etc. Due to inheritance of template instances (see Section 3.2) a single template instance may be used throughout the whole hierarchy of exhibition spaces.

Figure 5 shows a fragment of the X-VRML template that is used to generate the KML description. The template consists of KML and X-VRML commands (X-VRML commands are highlighted in bold). The fragment contains a loop that retrieves the names and assembles the URL addresses of all sub-spaces of the current space and inserts them in the KML code in appropriate places. The refresh interval is a parameter that can be set once for the whole template.

```
...
<For name="i" from="0" to="{sizeOf(@lSubSpaces)-1}">
  <Set name="currSpaceId" value="{@lSubSpaces[$i]}"/>
  <Set name="spPath" value="{@aPath}domain={@domain}..."/>
  <ARCO_AFProps afId="{$currSpaceId}" propName="SPACE_NAME"
    varName="spaceName"/>
  <NetworkLink>
    <name><Insert value="{@spaceName}"/></name>
    <description/>
    <visibility>1</visibility>
    <Url>
      <href><Insert value="{@spPath}"/></href>
      <refreshMode>onInterval</refreshMode>
      <refreshInterval>
        <Insert value="{@refreshInterval}"/>
      </refreshInterval>
    </Url>
  </NetworkLink>
</For>
...
```

Figure 5: *A fragment of X-VRML template for KML*

4.4. Schemes of interaction

Several schemes of user interaction can be implemented in the system depending on the templates used for generating the KML contents, objects' descriptions, and 2D/3D virtual exhibitions.

The basic form of interaction is when a user requests detailed information about a cultural object found in the Google Earth — either selected in the exhibition space hierarchy or on the map. The possible choices are: presentation of a 3D model of the object in the embedded browser, presentation of object's description in the embedded browser, presentation of a 3D gallery with focus set on the cultural object, and presentation of the object in the AR browser — launched as a separate application.

When a user selects the whole virtual exhibition — either in the exhibition hierarchy or on the map — the system may display the exhibition in 3D, in 2D or use it as contents for the external AR browser.

A user may also request geographical information about an object when browsing a 2D or 3D virtual gallery generated by ARCO. Using links accessible in the galleries a user may select presentation of any geographical information recorded for the object in the AMS metadata, e.g. indicate origin or current location of the object.

5. Examples of presentations

Several examples of the use of GeoARCO are presented in this section. In Figure 6, a simple presentation with three gallery spaces and four cultural objects is presented. The left side of the Google Earth application contains a panel that displays the exhibition hierarchy. On the map side, virtual exhibitions and objects are displayed. Exhibition spaces are marked with a house icon, while cultural objects with a flag icon. The cultural objects are positioned on the map in their places of origin. There is also an information window (displayed as a tooltip) for each cultural object with object's description, picture and links to the 2D/3D presentations of the object and a 3D gallery containing the object.

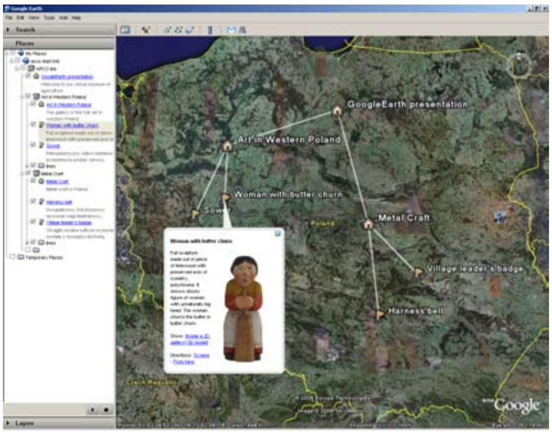


Figure 6: *Presentation of virtual exhibition spaces and objects on a map*

Figure 7 illustrates the case when a user requests detailed information about an object. In response, a 2D web page — dynamically generated based on a 2D presentation template — is displayed in the embedded browser. The page contains high quality picture of the object, full description, and selection of metadata elements. On this page a user has also access to all multimedia objects associated with the cultural object, e.g. movies, 3D models, pictures and sounds.

Figure 7: *Displaying information about a selected object*

In Figure 8, a presentation of a 3D model of the object is illustrated. The 3D model is displayed in the embedded browser on the right side. A user can freely manipulate the model to see it from all directions.

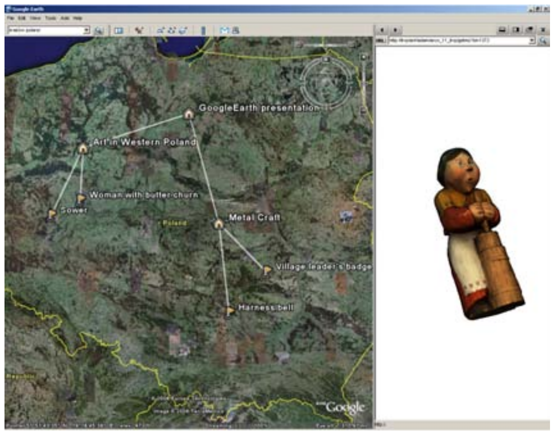


Figure 8: *Displaying a 3D model of a selected object*

In Figure 9, a complete 3D gallery containing two cultural objects is presented. The gallery is dynamically generated based on a 3D presentation template. The gallery is designed to allow users to navigate between the objects and manipulate their 3D models and display objects' metadata.

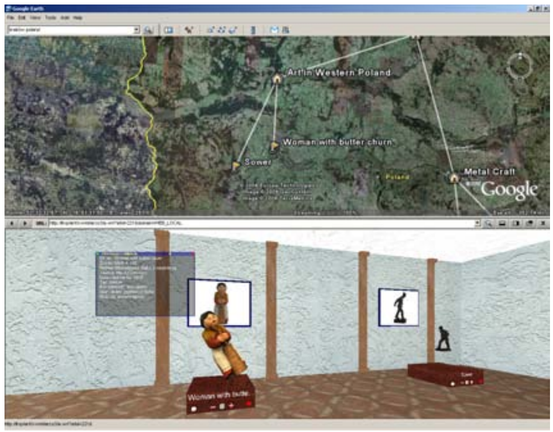


Figure 9: *Displaying a 3D virtual exhibition*

6. Conclusions and future works

The GeoARCO system presented in this paper enables museums to create high-quality, interactive virtual exhibitions that can be presented in the appropriate geographical context, thus considerably increasing their educational value. Users may browse the digital representation of the globe in the Google Earth system at the same time learning about the cultural heritage of particular regions. Users may also browse virtual galleries generated by the ARCO system and learn about origins or current locations of objects in the GIS presentation. By integrating the two advanced presentation platforms into a single interactive system, we have reached a new level of usability, that outperforms other solutions available today.

Important feature of the system is that 2D/3D cultural object representations and virtual galleries are built based on open standards and therefore can be used in other presentation modes (web, local kiosk, etc.). ARCO uses non-expensive hardware and software solutions. Even small museums with limited financial resources can afford the system. Also, the Google Earth platform is offered for free (basic version), making it accessible to all users.

We continue to develop the GeoARCO system. Future works include development of more advanced virtual exhibition templates enabling richer interaction with the Google Earth platform. Also, the KML templates may be extended to support geographical selection of objects and virtual galleries. Possibility to directly position objects on the Earth model would be desirable to simplify the use of the system. Based on the COLLADA description standard available in the newest version of the Google Earth system, we plan to build more advanced presentation elements that could be placed on the Earth model.

7. Acknowledgements

This work has been partially funded by the EU IST Framework 5 Programme, Action Line III.1.6 "Virtual Representations of Cultural and Scientific Objects" and by the Polish Ministry of Science and Higher Education (grant no 3T11C01428, 2005–2006). 3D models of cultural objects courtesy of the National Museum of Agriculture and Agricultural-Food Industry in Szreniawa, Poland.

References

[ARC] *ARCO — Augmented Representation of Cultural Objects*. http://www.arco-web.org/.

[BB] BEERS D., BERMAN N.: Zipusa: 58102 @ national geographic magazine. http://magma.nationalgeographic.com/ngm/0311/feature7/index.html.

[Col] COLLADA: *Collada Public Forum*. http://www.collada.org/.

[Elea] ELECTRONIC CULTURAL ATLAS INITIATIVE: *ECAI Iraq Home Page*. http://ecai.org/iraq/.

[Eleb] ELECTRONIC CULTURAL ATLAS INITIATIVE: *Electronic Cultural Atlas Initiative (ECAI)*. http://www.ecai.org/.

[Gooa] GOOGLE INC.: *Google 3D Werehouse*. http://sketchup.google.com/3dwarehouse.

[Goob] GOOGLE INC.: *Google Earth — Home*. http://earth.google.com/.

[Gooc] GOOGLE INC.: *Google Maps*. http://local.google.com/.

[Goo06] GOOGLE INC.: *Google Earth, KML 2.1 Specification*, 26 June 2006. http://earth.google.com/kml/.

[KB] KATO H., BILLINGHURST M.: *ARToolKit*. http://www.hitl.washington.edu/artoolkit/.

[Koe02] KOENEN R.: *Overview of the MPEG-4 Standard, ISO/IEC JTC1/SC29/WG11 N4668*, March 2002. http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm.

[Nat] NATIONAL AERONAUTICS AND SPACE ADMINISTRATION: *NASA World Wind*. http://worldwind.arc.nasa.gov/.

[New] NEW ZEALAND ARCHAEOLOGICAL ASSOCIATION: *Google Earth Archaeological Sites Australasia*. http://www.nzarchaeology.org/googleearthsites.htm.

[PWM*05] PATEL M., WHITE M., MOURKOUSSIS N., WALCZAK K., WOJCIECHOWSKI R., CHMIELEWSKI J.: Metadata requirements for digital museum environments. *Int. Journal on Digital Libraries, Special Issue on Digital Museum 3* (May 2005), 179–192.

[Sky] SKYLINE SOFTWARE SYSTEM: *Skyline Software System*. http://www.skylinesoft.com/.

[SKZ05] SITNIK R., KUJAWINSKA M., ZALUSKI W.: 3DMADMAC system: optical 3D shape, acquisition and processing path for VR applications. In *Proc. SPIE* (2005), vol. 5857, pp. 106–117.

[Tay] TAYLOR F.: *Google Earth Blog: Great Pyramid of Giza in 3D*. http://www.gearthblog.com/blog/archives/2006/05/great_pyramid_o.html.

[Uni] UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANISATION: *Map of World Heritage Properties*. http://whc.unesco.org/en/map.

[WC02] WALCZAK K., CELLARY W.: Building database applications of virtual reality with X-VRML. In *Proc. of the 7th Int. Conf. on 3D Web Technology (Web3D 2002)* (Tempe, Arizona, USA, 2002), pp. 111–120.

[WC03] WALCZAK K., CELLARY W.: X-VRML for advanced virtual reality applications. *IEEE Computer 36(3)* (2003), 89–92.

[WCW06] WALCZAK K., CELLARY W., WHITE M.: Virtual museum exhibitions. *IEEE Computer 39, Issue 3* (March 2006), 93–95.

[Web] WEB3D CONSORTIUM: *Extensible 3D (X3D) International Standard, ISO/IEC 19775:2004*. http://www.web3d.org/.

[Wil] WILDLIFE CONSERVATION SOCIETY AND NATIONAL GEOGRAPHIC SOCIETY'S EXPEDITIONS COUNCIL AND NATIONAL GEOGRAPHIC MAGAZINE: *MegaFlyover: Revealing the Last of the Wild*. http://www.megaflyover.org/.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

ICT investment considerations and their influence on the socio-economic impact of heritage sites

Jim McLoughlin, Jaime Kaminski and Babak Sodagar

Brighton Business School, University of Brighton, Brighton, UK.

Abstract

The following methodological model can be used as a platform for the study of the impact of information and communications technologies (ICT) at cultural heritage sites. The model has been developed through extensive, in-depth interviews with curators, directors and stakeholders at many cultural heritage sites across Europe. The underling strength of this model is its versatility. Although the model is oriented towards the investment in, and deployment of, ICT at heritage sites it is fundamentally about understanding the process of investment and so could be modified for many investment decisions.

Categories and Subject Descriptors (according to ACM CCS): K.6.1 Project and People Management: Strategic information systems planning.

1. Introduction

Heritage site managers and policy makers often ask the question: "What is the impact of a particular ICT deployment at a heritage site?" Usually, this kind of question is answered through an assessment of the usability of the technology concerned, or attempting to link increased visitation to the deployment. Although such analyses have their value they are simplistic. The impact of an ICT investment and deployment on heritage sites and their visitors is an incremental impact. That is to say it is an impact that occurs in addition to, and as part of the wider impact of the site. ICT does not exist in a vacuum divorced from the heritage system – ICT is part of the heritage system. The incremental impact of an ICT deployment cannot therefore, be viewed in isolation from the non-ICT impacts and outcomes associated with a particular heritage site. The success or failure of a particular ICT project is, more often than not, a function of factors outside of the realm of IT. Politics, design, and location amongst others play an important role in the success and failure of an ICT deployment. The success or failure of a project determines its socio-economic impact as much as the technology itself. It would be a gross simplification to think that technologies can be studied in isolation from these external factors.

2. Understanding the site

The holistic framework model (see Fig. 1) seeks to understand and conceptualise the dynamics of heritage sites. The framework consists of five elements: the cultural heritage site (CHS) impact context, the site mission and objectives, the site stakeholders, and the site management and decision making context, which all influence and contribute to the potential socio-economic impacts of a heritage site. This model provides a site context for the following model which is specifically oriented towards the deployment of ICT (see Fig. 2).

When studying the 'impact of technology' it becomes apparent that any analysis is meaningless without consideration of what makes each heritage site unique. Different sites have different strengths and weaknesses strong brands, exceptional collections, extensive financial resources, highly accessible locations, high footfalls, etc. Different sites also have different rationales and objectives for deployment of ICT. If the impact of ICT is divorced from these contextual factors then the result of a study will lose its meaning. This model allows those studying heritage sites to place them in the same conceptual framework.

2.1. The impact context

The impact context is interpreted broadly as the specific macro-contextual influences and micro-contextual (such as organisational) influences on a cultural heritage site. Macro contextual influences can include: macro-economic environment, policy context, legal framework, cultural context and values and technological context. The micro-contextual influences exist at two levels; the local environment and the site. These influences can include the local environment (economic, political, funding, demographic, legal, competition, infrastructure, etc) and the site (funding, ownership, governance structure, scale and location).

For heritage managers the impact context creates opportunities and threats for their organisations and can impose constraints on decision making. Most of these factors are beyond the direct control of cultural heritage managers, but nevertheless affect heritage site strategies and final impacts and outcomes. Furthermore, many of the factors are inter-related and so for example, local economy could affect heritage site funding or the policy context could affect the legal framework.

2.2. The macro context

Each site operates in a macro-national context (and wider European and global context). A number of influences from this context affect heritage sites, including:

- *Macro-economy*: The macro-economy (regional, national and international) affects, for instance, tax revenues, disposable income, and policy funding priorities. The macro-economy has a major influence on the heritage sector.
- *Policy context*: The macro-policy context is another important determinant for potential outcomes and impacts at heritage sites. Policy is fundamental to understanding impact; it influences heritage sites at multiple levels. It determines what gains funding and what does not, what is conserved and what it is not, it influences local authority policy, and it can also affect national legal structures which influence the heritage sector, etc. [MR04].
- *Cultural context and values*: The 'cultural context' and values of a society in supporting heritage, will in turn affect practical policy and funding priorities. For example, the cultural context helps define what is considered to be heritage. As such, definitions of what should be preserved can differ between countries. Furthermore, the definition of what constitutes 'heritage' is not static but dynamic. In the developed world the definition of heritage has broadened considerably in the later half of the twentieth century. It is important to acknowledge that as time passes the definition of heritage will continue to change according to different political aspirations, and the increasing input of communities and groups outside of the traditional field of experts.

2.3. The micro context

The micro context can be classified at two levels the local environment and the site:

- *The local environment*: Micro factors would include the local economy and local policy and political context. For example, numerous local authorities and governments have developed strategies, with accompanying funding, targeting heritage as a key element in regeneration programmes. In heritage sites with a strong orientation towards tourism, a principal element of a sites economic impact will depend on the total visitor experience which itself is dependent on numerous off-site factors (e.g. coordinated local tourism strategy, the presence of other visitor attractions, quality of facilities such as transport, restaurants, hotels, etc). It is rare for a heritage site to be immune to these factors.
- *Competition or complementarity*: The degree of competition or complementarity with other attractions can also influence impact. For example, a heritage site within a historic urban centre (such as Rome, Venice or Paris) could face competition from numerous alternative heritage attractions; however, the nucleation of heritage sites within a town or city can act as a stimulus to attract visitors. In such cases the visitors would be more likely to be interested in heritage tourism. Such situations have been given the label 'co-opetition'. Of course, the competition is not limited to other heritage sites, any attraction which could divert tourist money away from heritage represents potential competition, but the creation of a diverse tourist product offering is likely to be beneficial for attracting a more diverse range of visitors.

2.4. The site – organisational context

Organisational context is central to understanding impact. The impact of any site is heavily dependent on its location, quality, significance and the scale of the heritage site itself. As sites vary in their local, regional, national and global significance then so will their relative impacts. Some factors to consider include:

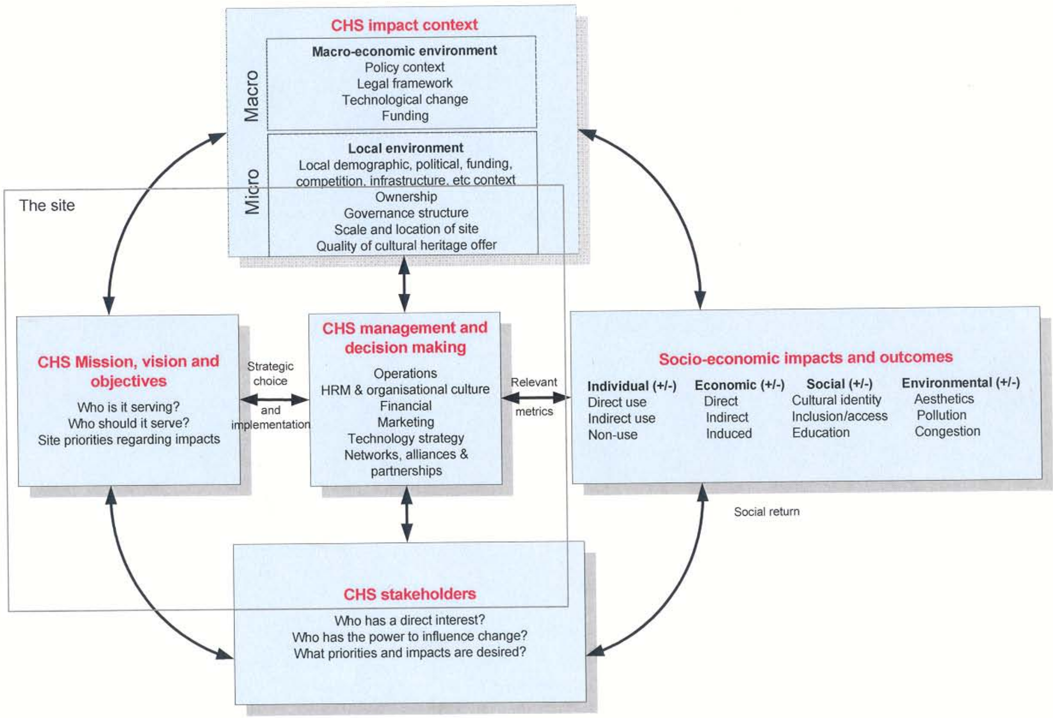


Figure 1: *A dynamic holistic impact model for cultural heritage sites [MSK06]*

- *Ownership*: The ownership of heritage sites is a key determinant of the impact that a site will have. Ownership influences funding sources, governance structures, objectives, etc. However, ownership of heritage sites is not static. For example, because cultural heritage sites can have high maintenance costs especially in countries with strictly enforced legislation regarding the upkeep of such sites. There is a tendency in such countries to see the movement of ownership from private to public hands.
- *Corporate governance*: Heritage sites can have a wide range of governance structures ranging from private and public, to not-for-profit and charities. Each of these will influence the impacts and outcomes of a heritage site. While it would be simplistic to assume that all sites under private ownership have a greater profit motivation than sites in public ownership there is a trend towards this scenario that cannot be ignored.
- *Location*: Location is paramount for the impact of a cultural heritage site. The location determines factors such as accessibility to transport networks, proximity to population centres links with other potential attractions. Surprisingly, location can be a dynamic entity. Although cultural heritage sites are fixed entities within the landscape or urban fabric the significance of the surrounding locality can change over time. A rundown part of an urban centre can become a popular tourist zone increasing the potential of the heritage sites within that area (such as the Gothic quarter Barri Gtic in Barcelona, preserved through neglect and now one of the principal tourist magnets in the city). Alternatively, the creation of new transport links such as low cost airline routes, or motorway and train-links can radically change the accessibility of a heritage site.
- *Quality of the cultural offer*: This exists at two levels. The significance of the site to society, and the quality of the 'visitor offer'.
- *Significance*: The significance and importance of a site is a difficult entity to define. Sites have significance at multiple levels such as local, aesthetic, regional, and national. Of course, as with so many elements of the dynamic impact context the signif-

icance of cultural heritage sites is not a static element, it can change over time. It can change because of changes in the political system, technology, etc. Even at a single point in time a site may hold alternative significance to different elements of the population this can determine who visits a particular site.

- *Quality of the visitor offer*: The quality of the visitor offer at a heritage site or experience can be determined by a number of factors such as the level of preservation, which lies outside the scope of the heritage site, however, site maintenance, level of restoration and visitor facilities tend to fall within the potential control of a site, finance depending, as can the actual or perceived authenticity of the site. Contemporary Western society is a consumer society. In this society the public have been exposed to progressively more sophisticated products, services and marketing, and as a direct result they have become much more sophisticated consumers. It is these same consumers who visit heritage sites and they will judge those sites accordingly. Facilities and services at heritage sites need to be at a standard commensurate with contemporary consumer ideals otherwise sites risk alienating many of their visitors.
- *Scale*: Scale can act as a guide to the potential impact of a cultural heritage site (although, no more than a guide). Larger sites have the potential to induce a greater impact than smaller sites, because of their ability to support a greater throughput of visitors, sustain larger potential capital costs, higher staff requirements and running costs. Of course, groups of smaller heritage sites can have a similar effect.

These factors have a strong influence on the site and feed into the management decision-making context. Furthermore, it is argued that contextual factors are immensely important determinants of the socio-economic impact of heritage sites. Placing a heritage site in context will guide what impacts that should be evaluated. For instance, there would be little point in doing a full, and often costly, economic impact analysis of a small local museum that was designed to serve the local community and foster local cultural identity. In such a context impact assessment should be aimed at issues of community integration and social inclusion, etc. In the dynamic model there is the potential for sites at the micro-contextual level to have some influence over the 'impact context'. Cultural heritage sites have a greater potential to influence and have an impact on the micro context compared to the macro context. Some heritage sites can make a (sometimes significant) contribution to the local economy through increased visitor numbers, capital expenditures, or brand value.

3. The technology impact context

The development of ICTs takes place outside of the cultural heritage sphere (usually in the commercial, scientific or military sectors) and gradually migrates to the heritage sphere. With new ICT hardware, software and their associated standards being developed continually, it is important to consider technological developments and how these might affect the visitor experience. Changes to the non-technology elements of the heritage site and its wider context can have wide ranging effects on the impact and outcomes of an ICT deployment. A number of factors affect the technological impact context, including:

- *Development of ICTs*: The ICT deployment in heritage sites exists within a wider 'ICT and technology' context. At the most fundamental level, what ICT is available is dictated by developments in the spheres of science, industry and commerce. Heritage sites do not have the resources or expertise to drive base-level change in ICT. But the availability of ICT is the principal determinant of what can be achieved.
- *Cost of technology*: Global economic forces have acted to drive down the price of ICT hardware and software. This contextual factor affects both heritage sites and their visitors. Lower costs have made ICT technology more accessible to heritage sites which tend to be characterised by limited finances (the increasing use of touch-screens, large LCD screens and solid state audio guides at heritage sites is an obvious manifestation of this). Moreover, ICT has become a commodity item in society. As more consumers have to opportunity to have increasingly sophisticated ICT in their homes, more people are becoming familiar with technology. Furthermore, many visitors will have access to technology in their workplaces. Visitors are therefore becoming increasingly familiar with ICT and so the accessibility has increased. This accessibility and familiarity can also lead to increased acceptance of technology. The visiting public are driving demand. Although, there is still a 'Digital Divide' within many European countries, which increases with the newly joined nations.
- *The acceptance of technology*: The acceptance of technology is determined by socio-economic factors. The widespread use of ICT is the result of complex interactions between economic forces and user needs. Acceptance of such technology is often dictated by the penetration of ICTs in society (the internet, digital TV, mobile phones, PDAs). The level of acceptance of technology is relevant to both the site visitors and the site interpreters.

- *Reliance on cutting-edge technology solutions*: Some

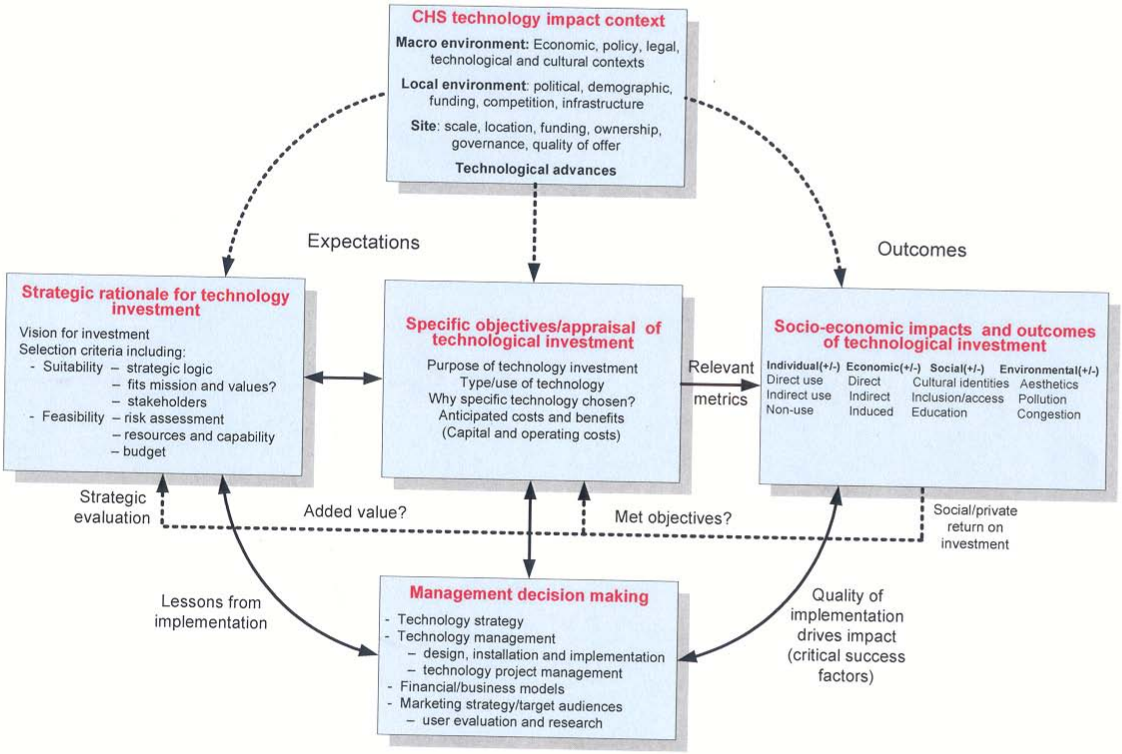


Figure 2: *A holistic investment framework that allows heritage site managers to conceptualise how ICT deployment influences socio-economic impact*

ICT technologies and standards are well-established (the Internet, PC hardware, HTML, XML, etc), but others are still in the process of gaining market acceptance. Heritage sites are not best placed to know which technologies and standards are likely to gain market acceptance, hence, why heritage has always been a late adopter of technology. Sites with potential ICT deployments that rely on cutting edge technologies/standards could run the risk of the technologies used failing to gain long-term market success, however, if successful these sites could have a market-leading advantage. Deploying technological solutions at the appropriate time is crucial.

3.1. Strategic rationale for technology investment

There has to be a strategic rationale for technology investment. This is usually closely linked to the mission and vision for the site. Strategy needs to underpin the management decision making at a heritage site. Two principal components are suitability and feasibility:

3.2. Vision

All investment decisions usually involve some intended innovation to enhance the cultural product offer. The vision is eventually a strategic view of where the site should be and what it should offer. Once this is clearly defined the exploration of the appropriate ICT for the vision can take place.

3.3. Suitability

- *Strategic logic*: There must be a strategic logic for the deployment of ICT. At its simplest a heritage site's strategy revolves around three questions: where is the site positioned now, where does it want to be positioned, and how will it achieve that goal. An ICT-based solution may, or may not, be the most effective use of resources for achieving that goal. There have been many examples of technology-led solutions that have been deployed at heritage sites for no other reason than the technology was available.

- *Site mission*: Another key question is does the particular use of ICT fit with the mission and values of the site? It is crucial that the deployment fits the mission and values of the site. For example, the type of ICT deployed at a site whose primary aim is education might differ from one where visitor numbers are required to support the revenue stream.
- *Stakeholders*: All investments involve opportunity costs. The potential funds that may be devoted to an ICT project can alternative uses. It is therefore essential that stakeholders support the deployment of resources.

3.4. Feasibility

- *Risk assessment*: The installation of ICT can hold considerable risk for heritage sites. For many it is an area beyond their traditional sphere of experience so they are reliant upon external sources of consultancy and services. A typical risk factor is cost outweighing the benefits.
- *Budget*: Sites have to consider if they have the budget for ICT installation and maintenance and/or the resources and capability to support such an installation.
- *Resources and capability*: The introduction of ICT requires numerous new skills. Heritage sites need to establish what resources and capabilities they have for such a deployment. Do they have any skills in house or will the entire project (or part of the project) need to be outsourced? Furthermore, ICT requires maintenance. Hardware which requires a high level of manual interaction such as touch-screens, trackballs, and keyboards all require upkeep. Purely electronic hardware such as processors, motherboard batteries, disk drives, can all fail. Bespoke software may have bugs. Sites have to allow for these contingencies and set aside resources at the outset for maintenance.

3.5. Management decision making

The management decision-making element is another key component that influences impact. This encompasses three components; technology management, the financial and business models, and the marketing strategy.

3.6. Technology strategy

Cultural heritage sites should have a continuous review of technology strategy (e.g. Web, audio-guides, booking systems, visualisation technology, etc) that can support the cultural offer.

3.7. Technology management

Technology management is a multi-faceted area:

- *Technology project management*: There are numerous considerations to be made when managing a technology project. For example does the project meet the vision. Is there a clear specification, as Soren [SOR05] notes "Clear objectives and values help curators take ownership of a project, and feel responsible for whether it succeeds or fails". It is necessary to liaise with external partners and with internal players (i.e. using human resource management for managing change). Not all heritage sites have the luxury of having full-time staff devoted to ICT management. Some have to share IT staff between sites or have staff who do IT-related tasks in addition to other jobs. These sites may have to purchase these skills from outside consultants. If the heritage site is for some reason unable or unwilling to maintain their ICT deployment then its impact may change from a positive to a negative. Furthermore, deploying ICT at a heritage site is not the end of the story. Information technology, as with all technology requires maintenance. Many sites do not have the skills to keep ICT projects running if the technology breaks down. This of course then requires external consultancy to fix any problems but, needs to be factored into the running costs of the original business and sustainability model. The following factors are also integral with technology management:
- *Management 'buy-in'*: Much work has been conducted in the commercial business sector that shows that the lack of senior management buy-in is one of the biggest reasons for the failure of technology projects. This is extremely important in the cultural heritage sector because there can still be reticence towards the use of information technology in what is still a sector with traditional origins. Without management buy-in projects could fail before deployment or could have insufficient resources for successful deployment, leading to negative impressions by visitors.
- *Leadership*: Closely related to the above is leadership. Leadership for an ICT deployment at a heritage site exists at two levels; the strategic leadership that drives the overall conceptualisation, and the IT project leadership that manages the actual day-to-day running of the project. Strong strategic and project leadership can greatly enhance its chances of success.
- *Design, installation and implementation*: When visitors come face-to-face with front-of-house ICT at

heritage sites their first impression is a function of the design, implementation and installation of the technology. The design of ICT applications is a complex area that is usually beyond the experience of heritage site personnel because so many different skill-sets are required (ICT development, graphic design, ergonomics, etc). As heritage sites have become more likely to deploy ICT to enhance the visitor experience this has created a market opportunity for organisations who design and install ICT solutions (and those who co-ordinate the various project specialists). Although, even today few enterprises can rely solely on the heritage sector for their business. Still heritage sites deploying ICT are now making a contribution to the business sector.

- *The quality of the implementation drives the potential impacts*: An exceptional use of technology can be let down by poor design, location, and implementation. Alternatively, lack of funding may result in poor design because shortcuts were made. This is important because considerable evidence points to cultural tourists as being increasingly sophisticated visitors. This does not imply that all visitors to heritage sites are classified as cultural tourists, but there is a tendency for museum and heritage site visitors to come from higher education backgrounds.

3.8. Financial and business models

In the past many heritage sites have been caught out by the lack of coherent, sustainable business models. Capital funds and grants have been devoted to projects but less consideration has been devoted to the sustainability of the project. There is evidence that this is slowly beginning to change many funding bodies now require evidence of sustainability and business planning before they grant capital funds to projects. In the UK funders such as the Heritage Lottery Fund and English Heritage now require sustainability plans for the projects they fund. There are numerous considerations for financial and business models, such as charging for specific exhibitions, developing exhibitions with the potential to tour and so gain extra revenue, or more imaginative models such as sharing development costs in return for a percentage of the revenue.

3.9. Marketing strategy and target audiences

- *Marketing strategy*: ICT deployments do not exist outside of a business system. If visitors are not motivated to go to the physical or virtual heritage site in the first place then the impact of the ICT deployments can be reduced. A significant investment in ICT might form the basis of a marketing campaign. At the British Museum the special exhibition the 'Mummy: the inside story' was accompanied by a strategic marketing campaign. This certainly increased the awareness and therefore had a considerable influence on the visitor numbers and so the scale of the impacts and outcomes.
- *User evaluation and research*: heritage sites have a long tradition of conducting research on their visitors to determine user satisfaction. Visitor surveys or interviews are and well understood by heritage sites. There is also considerable external consultancy available to sites (although, to date very few sites have targeted the incremental contribution to the user/visitor experience caused by the use of ICT). There is therefore a well-established mechanism that heritage sites can use to determine the socio-economic impact of technology at heritage sites. Furthermore, user evaluation can be used to support marketing research.

3.10. Specific objectives and appraisal of the technology investment

The purpose of technology investment is often key for understanding the impact of ICT. ICT investment reflects cultural product innovation and can provide a basis for a new offer. There can be a wide range of reasons for the deployment of visitor-facing ICT at heritage sites. These can include:enhancing the users experience, increasing visitor numbers, increasing accessibility, enhancing educational impact, or some combination of the above.

A key question that sites often want answered is "has the investment achieved this aim?" The objectives of a project are key to determining what impacts should be assessed.

- *Type/use of technology*: The purpose for a technology investment is a key determinant for why a specific technology is chosen. This of course is tempered by the anticipated costs and benefits of such a deployment. The type of technology chosen is key for impact assessment. Different technologies have different potential for impacts and outcomes. Technology that is connected to the internet may have a greater impact because of the potential for access to a larger number of people. Visualisations at heritage sites may have a considerable impact to the visitors, but this may not be translated to a broader impact because of the localised nature of the impact.
- *Anticipated costs and benefits*: This is the essence of appraisal. The initial capital cost outlay can be estimated as can the potential social returns and benefits. The anticipated costs may be assessed through the use of Return on investment (ROI), and Net

Present Value (NPV) calculations. It is essential to consider both the capital and operating costs for a deployment. These assessments can then be compared to the potential anticipated benefits that the use of ICT may entail. Once a project is running the impact measures can be used to provide data on the actual return.

4. Socio-economic impact of technology investment

In this model the measurement of socio-economic impacts is not just something that is necessary to fulfil funding obligations, but is an essential part of the management decision-making process of a heritage site. The measurement of impacts is key to validating the heritage site strategy. Socio-economic impact embraces many possible impact dimensions (e.g. economic, individual, social, environmental, etc). Within each of these dimensions there are a number of possible methodologies which can be employed to identify and measure impact, each method having advantages and disadvantages [MSK06].

5. Conclusions

The above models highlight the limitations of assuming a simplistic relationship between deploying technology and its impact. It is apparent that a multiplicity of factors influence social and economic impact simultaneously with any technology impacts. Deploying ICT is therefore no guarantee of achieving the goals of a site or improving the deficiencies of a site. However, if there is a strategic rationale for technology investment then there is greater potential for positive impacts and outcomes. The break down of the model into elements allows users to conceptualise the process of investment. This way of thinking could be called 'heritage systems analysis'. This is to say a consistent theoretical model for heritage sites that allows the internal and external factors that influence impact to be conceptualised. If the heritage sector were to understand how various components of the system are interlinked and affect impacts and outcomes then this could become the basis for understanding impact. In this context understanding impact becomes the basis for positively influencing impact. The underling strength of this model is its versatility. Although the model is oriented towards the investment in, and deployment of, ICT at heritage sites it could be modified for many investment decisions.

6. Acknowledgements

EPOCH's socio-economic impact research has been supported by the European Commission's Framework 6 programme (IST-2002-507382).

References

[MR04] MIGNOSA A., RIZZO R.: Heritage and information. In *Interdisciplinarity or the best of both worlds – the grand challenge for cultural heritage informatics in the 21st century, Selected papers from VAST 2004* (2004).

[MSK06] MCLOUGHLIN J., SODAGAR B., KAMINSKI J.: Dynamic socio-economic impact: a holistic analytical framework for cultural heritage sites. In *Heritage Impact 2005: Proceedings of the first international symposium on the socio-economic impact of cultural heritage* (Budapest, 2006), Archaeolingua, pp. 43–57.

[SOR05] SOREN B. J.: Best practices in creating quality online experiences for museum users. *Museum management and curatorship 20, 131-148 20* (2005), 131–148.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Watermarking and Digital Rights Management - A Pilot DRM System Implementation and Technical Guidelines to Cultural Digitization Projects

D. K. Tsolis[1] S. N. Nikolopoulos[1] E. G. Karatzas[1] S. A. Sioutas[1] E. A. Hondrou[1] A. I. Mouriki[1] Ai. A. Georgiadou[1] and T. S. Papatheodorou[1]

Department of Computer Engineering and Informatics, University of Patras

Abstract

The issue addressed in this paper is at first a brief presentation of the Technical Guidelines for IPR protection and management applied to Greek cultural digitization projects. Secondly, the work focuses on the analysis and implementation of a typical Digital Rights Management System for organizations and projects aiming at the digitization and exploitation of cultural content. Both Technical Guidelines and the DRMS are setting a solid framework for providing answers to a crucial and complex issue, the issue of the protection and management of intellectual property rights for analog and digital content.

Categories and Subject Descriptors (according to ACM CCS): H.4.0 [Information Systems Applications]: General

1. Introduction

Greece has already recognized the great importance of digitization for Cultural Heritage. The salvation, the long term preservation and exploitation of cultural resources of any type are strongly related to digitization (a process that includes, the creation of digital surrogates, the structuring of the digital resources in repositories and the content management aiming at efficient exploitation through added value services). In addition, long term digital preservation of digital memory is considered an emerging technological and policy issue. Both issues of digitization and long term digital preservation are considered as the basis of promoting Greek Culture through the Internet, and through interactive and multimedia applications.

Nevertheless, before the process of digitization to take place, there is one issue, which is constantly emerging, setting barriers not only to digitization itself but also to further exploitation of the digital content through the Internet and E-commerce applications, the issue of copyright protection and management. Dealing with this issue, a wide range of activity has been initiated focusing on setting a functional legal framework, supporting organizations through Technical Guidelines and experimenting with rising technologies like the Digital Rights Management Systems. The results of these activities are presented through this paper, shedding light to an issue of worldwide interest. Based on the research activity carried out at the High Performance Information Systems Laboratory of the University of Patras, a pilot information system is also being presented, which supports a cultural organization to protect and manage rights for digital content through watermarking. The system was initially designed and implemented for the Hellenic Ministry of Culture. The objective is to propose the main features, functional and technical requirements and innovative subsystems of a typical Digital Rights Management System for cultural organizations and the results to be adopted by cultural public or private organizations. The system utilizes innovative watermarking algorithms, digital image libraries, interoperable e-licensing mechanisms, unique identification subsystems so as to support digital rights management and protection functions.

2. Technical Guidelines

The Technical Guidelines entitled "Practical guide-handbook for the protection and management of Intellectual Property Rights of digital cultural content and a digital rights management system based on international metadata standards", was based on an extensive research study and a pilot implementation on relevant issues. The aim was to support cultural organizations to characterize content to be digitized, clear the intellectual property and relevant rights. In parallel, gives advice on who are the key players, what are the collective societies and clearing houses in Greece, when and how to get in contact with them so as to clear rights and comply to the current legislative framework.

In addition, the cultural organizations are introduced to the current technological solutions for copyright protection and management such as watermarking, encryption, metadata and Digital Rights Management systems in general. Systems, software and hardware solutions, commercial applications which are focusing on the copyright protection and management issue are being presented and analyzed. The Technical Guidelines are not aiming at evaluating the software and hardware solutions but are giving advice to the cultural organizations of how to choose a solution that fits their needs amongst the existing ones.

The Technical Guidelines main scope is to provide a direct and effective method of navigating through the subsequent issues of the problem. The issues are not exhaustively analyzed and described. At first the issues are briefly presented and then the reader is directed to more detailed sources of information (if desired). In this way organizations are able to follow a step by step approach for resolving their rights clearance issues and are at the same time supported by to-do lists and tips. The process results at a clear set of rights for the digitized content, with special terms of use for Internet and E-commerce applications with relevant electronic licenses produced between the cultural organizations and technological providers.

The first chapter of the Guidelines is a practical guide to national and international legislation. A clear view of the restrictions for the use of digitized content is being given to cultural organizations, especially for the content exploited through the Internet. Furthermore, the technological means and the functional requirements necessary to support cultural organizations to protect and manage the intellectual property rights of the digital content.

In the framework of the Guidelines, several flow diagrams for the process of clearing rights are being given, which when followed by the cultural organizations (assisted by certain choices), they manage to clear the rights of the digital content, to define all the details regarding the terms of use, restrictions of use and embed watermarks for the protection of the copyright.

At present time, there are more than 300 on-going digitization projects in Greece of about a total of 120 million Euros cost, which are developing the Greek digital cultural content, the services and accompanying tools and are obliged to conform to the aforementioned Technical Guidelines. The document is only in Greek and is distributed in electronic format by the Greek Information Society Secretariat. These projects are core digitization projects aiming mainly at digitizing, disseminating and promoting collections of various historical periods of Hellenic Culture. The collections that are being digitized are including various types of content, images, artefacts, sites, buildings, archives, music, cinema and movie libraries, etc. The use of Internet for the promotion of the digital content is a very important work package of the projects, as well as the dissemination of cultural objects in other languages (English, French, German, etc). The results of the projects will be the mass digitization and mass promotion of the Hellenic Cultural Content in a European and World wide level. In addition, the content industry in Greece is being supported and cultural organizations have an important opportunity to digitize and safeguard their collections and artefacts.

3. A Typical DRMS for Digitization and IPR Protection

In this section and based on research activity carried out at the High Performance Information Systems Laboratory of the University of Patras (http://www.hpclab.ceid.upatras.gr) [DT1], a pilot information system is being presented, which supports a cultural organization to protect and manage rights for digital content. The system was initially designed and implemented for the Hellenic Ministry of Culture and the results are partly used in the aforementioned Technical Guidelines. The work was not aiming at promoting an already implemented DRMS. The objective was to shed light on the main features, functional and technical requirements and subsystems of a typical Digital Rights Management System for cultural organizations and the results to be adopted to the Technical Guidelines.

3.1. What is Digital Rights Management?

The term Digital Rights Management (DRM) was introduced in the late 1990s [CSTB99]. When content is created (information), a control to a set of rights to that content is inherited to the owner - allowing browsing, editing, printing, executing, copying etc. Traditionally, those rights have accrued from three sources:

- Legal. Rights that someone acquires either automatically under law or by some legal procedure (such as applying for a patent).
- Transactional. Rights that someone gets or gives up by trading them, such as buying a book or selling a manuscript to a publisher.
- Implicit. Rights defined by the medium that the information is in.

The most important matter about DRM is that the first two sources of rights haven't changed much with the advent of technologies such as the Internet, cell phones and MP3 files. Various parties have called for a complete replacement of the Intellectual Property (IP) law without correspodence.

The legislators have responded to new technologies by creating Directives and Acts. Transactions have remained the same regardless of the fact that they can be performed over the Internet. What is different is the implicit nature of rights when applied to traditional media. The Internet has made these implicit rights explicit. This engenders problems as well as opportunities for content providers as well as consumers [NK02].

Digital Rights Management refers to digitally controlling and managing rights for analog or digital content. The need for control and management has increased now that digital network technologies have taken away the implicit control that content owners get with legacy media.

3.2. The Typical DRM System - an Overview

The DRMS's main objectives are [BR02]:

- To provide an appropriate information infrastructure, especially focusing on cultural digital content (digital images) and its special characteristics. The services implemented range from typical e-commerce applications (electronic catalogs and shopping kart) to advanced services such as searching for images based on the image content and unauthorized content use detection.
- To protect the copyright of the digital images though robust watermarking techniques. Multi-bit watermarks are embedded to the digital images which are commercially exploited and delivered to the buyers.
- To support the digital rights management process for the cultural content and for the transactions taking place.
- To provide an effective mechanism for tracking down improper use of digital images which are owned by the cultural organization.

The general system architecture and its main components are the following:

- The Digitization layer.
- The Digital Image Library.
- The copyright protection subsystem, which protects digital content with watermarking techniques and provides for digital rights management.
- The E-Commerce applications.

The following information includes a detailed presentation of the main features and characteristics of the subsystems which the typical DRMS consists of. These systems and applications are:

- **Unique identification system**. One of the key challenges in the move from physical to electronic distribution of content is the rapid evolution of a set of common technologies and procedures to identify and manage pieces of digital content. A widely implemented and well understood approach to naming digital objects is essential if we are to see the development of services that will enable content providers to grow and prosper in an era of increasingly sophisticated computer networking. The International DOI Foundation (IDF) [DOI] was established in 1998 to address this challenge, assuming a leadership role in the development of a framework of infrastructure, policies and procedures to support the identification needs of providers of intellectual property in the multinational, multi-community environment of the network. The IDF has developed, and continues to evolve, a fully implemented solution to this challenge: the DOI System, using the Digital Object Identifier (DOI), an "actionable identifier" for intellectual property on the Internet. The DOI is now widely implemented by hundreds of organisations through millions of identified objects. The implemented DRMS is utilizing the DOIs infrastructure for the unique identification of the elements of the digital content.
- **Digital image library**. The design and implementation of the Digital Image Library is required for further development of the DRMS. The Digital Image Library is consisting of the Digital Image Library and the Metadata sets which are described in detail. The efficient management of digital images is based on an advanced database system. The Digital Image Library is designed and developed in accordance with metadata sets described in the following paragraph. The metadata sets are incorporated through tables, fields, triggers and views in the Database. The need for adopting international metadata standards is profound, especially for applications aiming at cultural content exchange. The DIG 35 Specification "Metadata for Digital Images" [DIG35], holds a very important role in the selection of fields and tables, regarding the digital images metadata. This metadata standard is already being widely used in simple end-user devices and even to worldwide networks. The database structure has also a special focus on metadata for the Intellectual Property Rights management.
- **Copyright protection subsystem**. The copyright protection subsystem is an intermediate layer between the e-commerce applications and the digital image library. Its main function is to protect the copyright of the digital images stored and exploited by the DRMS. Using a simplified view of the subsystem, it is considered as a black box which takes the original digital images as an input and produces the watermarked images. The whole process is automated and whenever a new original image is stored to the digital library the watermarked surrogates are created which carry the copyright owner id and other information used for copy control, digital signature, unauthorized use tracking and transaction management.

Watermarking Algorithm. Watermarking principles are mainly used whenever copyright protection of digital content is required and the cover-data is available to parties who are aware of the existence of the hidden data and may have an interest removing it [SK00]. In this framework the most popular and demanding application of watermarking is to give proof of ownership of digital data by embedding copyright statements.

For this kind of application the embedded information should be robust against manipulations that may attempt to remove it [PW02]. Many watermarking schemes show weaknesses in a number of attacks and specifically those causing desynchronization which is a very efficient tool against most marking techniques [CV03]. This leads to the suggestion that detection, rather than embedding, is the core problem of digital watermarking. According to the above the first most important step towards the implementation of the watermarking algorithm is the selection and evaluation of the watermarking method. The method chosen is mainly based on the further elaboration of the MCWG (The Multimedia Coding and Watermarking Group, http://www.mcwg.gr) watermarking tool, focusing on constructing a more efficient detection mechanism, resulting to a more robust watermarking technique. The core of the MCWG tool is a transform domain technique that is based on the use of the Subband DCT transform. The marking formula is the same well known multiplicative rule used in the large majority of the existing literature. The proposed watermarking method was tested particularly with digital images provided by the Hellenic Ministry of Culture and fine-tuned in accordance with the produced results. In addition, certain actions were taken for the further development of the method so as to incorporate multi-file support, monochrome and colour images and multidimensional digital images.

- **Effective licensing mechanisms**. The difference between purchasing and licensing is very important. Purchasing a copy of a work is the dominant transaction model of the copyright legislative framework for more than 200 years. The purchasing process, involves the exclusive transfer of ownership rights from the creator to the buyer. The copyright directive allows, for certain goods, the buyer to loan, rent or resell the purchased copy. Licensing, on the other hand, consists of a restricted transfer of rights for the work for certain uses under defined and declared terms and conditions. The licenses are expressed through contracts between the interested parties. The contract includes a wide range of terms and conditions. Licensing is widely used also for information previously embedded in cultural heritage objects. Although, certain license types might have advantages, their use as an information distribution model, provokes scepticism, especially pertaining to the restriction of public access to digital content of high educational value. Licensing negotiations tend to be time consuming. The adoption of an e-licensing mechanism for the typical DRMS is necessary and is implemented through technological standards which facilitate interoperable and direct licensing modes. The license produced by the DRMS is coded in XrML (eXtensible Rights Markup Language). Whenever an element of the digital content is being transacted (distributed, purchased, etc.) a digital license is being produced by the DRMS and distributed to the interested parties.
- **E-commerce applications**. The last layer of the DRM system's architecture is the one that provides all the e-commerce applications and services. These applications aim at establishing new standards in the field of e-commerce mainly in the digital cultural content sector. The most important components are:

  Definition of a standardized pricing policy specifically for the digital images of the Cultural Heritage.

  Flexible on-line license agreement which defines restrictions of use and rights for personal use.

  Design and implementation of an on-line shop based on the Digital Image Library with an advanced on-line catalog.

  Methods of secure commercial transactions using watermarking technologies.

  The standardized pricing policy for digital images purchased through the web is promoting a flexible user license agreement ("signed" on-line). The users that adhere to the terms of the license have the right to reproduce a digital image, to use the digital image in web sites, CD-ROMs, to edit the image and create original works, but do not have the exclusive rights to resell the digital image or indirectly gain profit based on the digital image.

3.3. Case Study: Hellenic Ministry of Culture

The aforementioned DRMS was experimentaly installed to the Hellenic Ministry of Culture. The system is being used by personnel of three authoritative agencies of the Ministry. In everyday use the system supports:

- The distributed insertion and management of surrogates in the selected organizations.
- Automated and imperceptible embedment of watermarks for the copyright protection of the digital content.
- The insertion of twenty high quality digital images per day and per organization.
- Once a month and when the Ministry's network is not overloaded, these digital images are safely transferred through the network to the central database server.
- Low quality, watermarked copies are automatically presented through the Ministry's Web Site.

The above scenario is based on the assumption that a trained user in every organization will be using the system five hours per day.

4. Conclusions

The Technical Guidelines was based on an extensive research study and a pilot implementation on relevant issues. The aim is to support cultural organizations to clear rights, to comply with the current legislative framework and protect and manage the copyright of digital content.

A typical DRMS is an integrated information system able to carry out numerous functions on behalf of cultural organizations, pertaining to the protection and management of digital cultural content, and its copyright. The DRMS provides:

- Services for creation, management and long term preservation of cultural content.
- Digital management of rights and the copyright of the content.
- Copyright proof and protection of the digital content through technical means (e.g. watermarking).
- E-Licensing mechanism for direct and effective licensing for digital content.
- Added value services for the end user (e.g. e-commerce applications).

The technologies used include digitization, digital libraries, digital object identification, metadata, encryption, watermarking, XrML, e-commerce applications and services etc. The effective combination, customization and integration of these technologies into an information system support the effective protection and management of the copyright of the digital cultural content.

5. References

[BR02] Bill Rosenblatt, Bill Trippe and Stephen Mooney, Digital Rights Management - Business and Technology, Professional Mindware, M and T Books, New York, 2002.

[CV2003] V. Cappellini, A. Piva, D. Dawson, D. Tsolis, "Technological Solutions for Copyright Protection and Management of Cultural Heritage: Best Practices and Guidelines", Workshop on Digitization of Cultural Content, 27-28 June 2003, Corfu, Greece.

[CSTB99] Computer Science and Telecommunications Board, National Research Council. (1999). The Digital Dilemma: Intellectual Property in the Information Age (pp. 2-3). Washington: National Academy Press.

[DT01] D. K. Tsolis, G. K. Tsolis and T. S. Papatheodorou, "A watermarking environment and a metadata digital image repository for the protection and management of digital images of the Hellenic Cultural Heritage", International Conference on Image Processing, Thessalonica, Greece, 2001.

[DIG35] DIG35 Specification - Metadata for Digital Images. Version 1.0. (2000). Digital Imaging Group.

[DOI] Digital Object Identifier, http://www.doi.org

[NK02] Naomi Korn, Guide to Intellectual Property Rights and Other Legal Issues Version 1.0, Minerva Project, http://www.minervaeurope.org/publications/guideipr.htm.

[PW02] Peter Wayner, Disapperaring Cryptography, Morgan Kaufmann, 2002.

[SK00] Stefan Katzenbeisser, Fabien Petitcolas, Information Hiding, Artech House, 2000.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Deconstructing the VR - Data Transparency, Quantified Uncertainty and Reliability of 3D Models

S. Hermon[1], J. Nikodem[1 2] and C. Perlingieri[1,3]

[1] Vast-Lab, PIN, University of Florence, Italy
[2] University of Bielsko - Biała, Poland
[3] University of Naples, "l'Orientale", Italy

Abstract

The paper discusses two key concepts required for the use of Virtual Reality and 3D modelling as a research tool for the humanities: data transparency -what is the type and nature of the archaeological/historical/ethnographical material on which the 3D model is based, and reliability - how the user can scientifically analyse the model. In this article, we will present a solution to these issues based on concepts deriving from fuzzy logic and fuzzy sets. Taking into consideration the "real nature " of humanities data, more often fuzzy than crisp, a different logic (fuzzy logic) should be applied when attempting to reconstruct a past reality. This will enable a quantifiable visualization of possible scenarios, otherwise discarded in traditional representations. Each scenario is accompanied with a "reliability index", estimated by the researcher according to his/her certainty on the existence of the modelled part and the "importance" of each component of the model. This approach will allow the user to reconstruct the "cognitive process" and the step-by-step "decision making" of the researcher that built the 3D model, and to open the model to a scientific analysis from a humanities point of view.

Categories and Subject Descriptors (according to ACM CCS): I.6.5 [Model Development]:

1. Introduction

Virtual Reality and 3D modelling applications to humanities and social sciences have a already long history of more than two decades, traceable to the early eighties of the last century. Though mostly applied to education or dissemination to the general public, theoretical and methodological issues regarding these kinds of applications to humanities have been raised and largely discussed in the past 15 years [Rei89], [RS89], [Sim97], [FS97], [BFS00], [Nic02]. The potential of VR as a communication tool has been highlighted by Ryan [Rya96], while accuracy of final product [Kan00], [FAF*00] and the need for data transparency [For00] are still commonly discussed issues.

Despite the increasing number of VR and 3D models of humanities data produced and reported [Bur02], [DS03], [NH06], most papers report issues related to communicating Cultural Heritage to the public or related technological aspects and improvements [ACN03], [CCSN04], rather successful stories related to the use of VR as a scientific tool in the humanities, a "pitfall" already pointed out a decade ago [Rya96], and since then rarely raised into the debate [Vat02], [BP04], [HND05]. An emerging legitimate question, after almost a quarter of century of work, is: "why should humanities and social sciences invest effort in VR and 3D modelling?". A more restricted version of the previous question is "how can humanities and social sciences exploit for research VR and 3D modelling applications?". Even though this paper is focusing on the use of VR and 3D modelling in the archaeological research, it is our belief that same arguments raised here are applicable to other fields of humanities and social sciences as well.

The article is proposing a methodology, based on the foundations of fuzzy sets theory, to evaluate the reliability of 3D reconstructions and thus open them to "archaeological scientific criticism". Its main steps are identification of basic components of the model, assignment of an index reflecting the reliability and importance to each one of them, computation of an overall reliability index of the entire model and an evaluation of the reliabilities of other possible reconstructions.

2. VR and 3D modelling and archaeology

Computer-based reconstruction of ancient monuments are increasingly confirmed to be extraordinary tools for representing the past reality, being archaeology a destructive technique employed on largely damaged contexts. One can visually express alpha-numeric data and graphically express thoughts and ideas, translating "...empirical phenomena into geometric language..." [FNRB02]. This aspect was well sinthesized by Niccolucci among others [Nic02]: "...since interpretation, explanation and communication involve reasoning, Virtual Archaeology can provide virtual creations to organize and synthesize known facts, showing them with greater clarity to others or to one's "inner eye", or virtual substitutes of physical objects...". The positive relationship between visualization tools and the way we perceive information has been demonstrated also by numerous researches in education, cognitive psychology and related disciplines [EFH76].

In order to accept as scientifically valid a 3D model, its resources and the criteria on which they are based should be explicitly presented [All98], and when possible, alternate reconstructions should be made available to the user [RR97]. Moreover, a 3D model can be considered scientifically reliable (the term "scientific" implying the Galilean approach, that can be repeated with the same result, beyond experimental errors) when it can be de-constructed and its data made transparent for a critical evaluation.

A typical 3D model (or VR) of an archaeological subject is build upon data of various sources (historical, linguistic, archaeological or history of art researches, just to name a few). Comparative and ethnographic studies, as well as the accumulated knowledge of the researcher are other factors that play a determinant role in the final shape of the 3D product. It is clear that without a clear and detailed description of the data and their sources a 3D model cannot be considered scientifically reliable. However, by merely presenting this metadata is only one step forward towards the use of a 3D model as a source of scientific information.

Incompleteness is a concept implicit in archaeological studies. Archaeology, and any other discipline that integrates archaeological data aiming at creating a virtual model, de facto do not provide enough data to establish a scientifically unquestionable model. To a certain point, all reconstructions remain "speculative". We should therefore consider as scientifically valid a model that allows us to quantify the "degree of speculation" and express it through established rules.

In other words, an important addition for turning a 3D model into a valuable source of scientific information is the possibility to evaluate the decision - making process of the researcher building the model, particularly in cases when (s)he is faced with several reconstruction options. This is mostly pertinent when data sets are not crisp (the window might have been square, or circular, or possibly elongated, and located at 2, 3 or 5 m from ground). Commonly, by arbitrarily deciding which window to model, and at which position in space, all other possible alternatives are discarded and disappear as soon as the model is completed.

It is proposed to apply concepts of fuzzy logic whenever fuzziness is identified in the data sets used for the 3D model, in order to: a. quantify reliability and uncertainty of reconstruction and b. introduce the concept of "fuzzy reasoning" to visualize "possible existences", each with its reliability index [HN03], [HND05]. In this way, not only the decision-making process becomes more transparent, but the user is provided with alternative choices for the 3D reconstructed elements, each with an index reflecting its possible existence. We will exemplify how these concepts were applied by us to the 3D model of a house from the archaeological site of Qalqal Asba.

3. Description of the case - study

The 3D model used as a case-study in this paper was created by a team of the University "l'Orientale" of Naples, as part of a wide research project at Aksum/Bieta Giyorgis, Ethiopia [PMP]. It represents the main building structure at the site of Qalqal Asba, an Aksumite site dated between 350 and 550 A.D. [BDF*02].

Translating site information into visual images required a complete and throughout analysis of all site data, in particular a throughout spatial analysis of archaeological features [PMP], [MPPF03]. Moreover, the virtual reconstruction of the structure (Figure 1) compelled the 3D model's authors to fill gaps and inconsistencies in available data and thus to critically analyse their archaeological data and cope with new questions arisen during the process of 3D modelling.

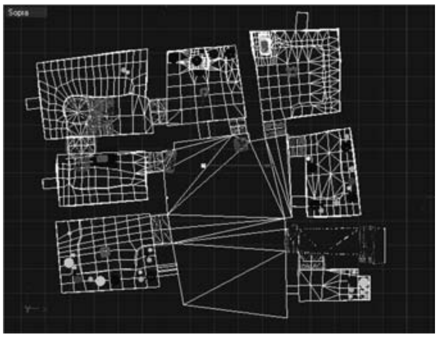


Figure 1: *General plan of the structure*

The Aksumite construction technique, with walls built of loose undressed stones, caused the almost complete collapse of the structure; only 2-3 lower wall courses were preserved. The building has a "U" plan, with a central courtyard open on the west side and nine rooms, all with access from the courtyard except for one, which had only internal access from the rooms beside.

Rooms are square or sub-rectangular, with stone benches along the walls. Occasionally, the roof was sustained by a wood pillar (not found) on a stone base. The characteristics of the rooms and an inventory of material culture allowed an estimation of their original function (Figure 2).

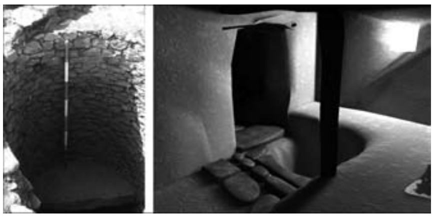


Figure 2: *The store-room: architectonic details and virtual reconstruction*

Remains of roof and second floor were almost impossible to recognize; one of the reasons for this is that these features of Aksumite buildings were constructed by layers of small stones and mud supported by a lattice network of wooden beams and smaller branches. Missing information from archaeology was counterbalanced by the ethnographic research, which provided information on the traditional architecture techniques and materials, spatial organization, and remains of material culture (Figure 3). Visits at a number of abandoned houses gave insights into the residuals, time and dynamics of collapse and depositional processes [BDF*02].

Figure 3: *Etnographic survey and the 3D model*

4. Evaluation criteria

The evaluation criteria of the reliability of the 3D model were based on a combination of two indices: importance and reliability, for each of the 3D model's features. Since the aim of the discussed 3D model was to understand the original architecture of the main building, the values of the above mentioned indices reflect the contribution of each analyzed component to its general shape.

Architectural units, such as walls, roofs or pavements were important for a general representation of the structure, while location of doors and windows, the interior design and furniture were significant in understanding aspects of social organisation, labour division and function of particular features. Moreover, by choosing to model particular decorations for doors, or a particular type of roof, style and cultural aspects of the ancient inhabitants of the site were evidenced. These were grouped under the term "conceptual categories" (see below) in order to express concepts that could contribute to the vision of the whole structure and as a sort of "bonding agent" for all the above categories.

Thus, components were divided into four main groups: components that directly affect the frame of the structure (shape and location of foundation, walls, roof, etc.), components with a secondary architectural role (shape of windows, doors, etc.), components of interior design (furniture and other material culture objects, such as grinding stones, fireplaces, etc.) and finally components that depict conceptual ideas (division and use of social space, etc.).

A reliability and an importance indices defined by values deriving from fuzzy logic concepts - that is belonging to the whole [0,1] interval - were assigned to each reconstructed component (Figure 4). The factors taken into consideration when assigning the reliability index were based on the source of data: primary (excavation), secondary (analogies), and then assumptions made on scientific deductions or pure imagination. The importance index was estimated according to the aim of the model, as emphasized above. Translating text concepts to numerical values, we obtained value intervals as follows: very important = value 1-0,8; significant = 0,7-0,5; scarcely important = 0,4-0,2; insignificant = 0,1-0.

5. Basic concepts of fuzzy logic and fuzzy sets theory

Fuzzy logic [Zad65], [Zim84], [Nov89], [Kos93], [MF93] parts from the observation that the world is composed by shades of gray, rather than sets of blacks or whites. The argument of the Sorites Paradox (what is a heap) may be used as an example for the recognition of vague concepts [FIS00], as opposed to Boolean ones (fuzzy against crisp). An important advantage of fuzzy logic is that it operates in the grey field of the uncertainty and ambiguity; using fuzzy sets one can handle imprecision. Moreover, fuzzy logic aims to extend ordinary deductive methods to situations in which the information available may be only partially or approximately true.

Since its introduction some half a century ago, fuzzy logic has been applied in natural, social and human studies, engineering and computer sciences, covering practically most fields of modern research, a large part of our modern daily life depending on fuzzy logic based systems and machinery, and the list is still long. Fuzzy logic was applied in social sciences as well [Rag00], anthropology [Rui99], databases [Cho03], economy [Dol94], religious studies [Rap93], vegetation science [Hal97], [Mor93], [Muc97], [OLGE98], [Rob86], [Rob89], [Sat96], or geography [Fis00], [HSB93], to name just a few.

While fuzzy logic is mostly used to solve problems concerning "de-fuzzification" of processes or prediction of models with a fuzzy nature, the archaeological reality, where a definite "true" or "false" will be possible only with the invention of Well's time machine (most probably never!), is intrinsically fuzzy and therefore should be treated as such. Since many of our concepts in archaeology are best defined as vague (fuzzy), and the archaeological reasoning is more often based on approximations and evaluations rather than on crisp affirmations, fuzzy logic offers an alternative methodological framework and the fuzzy set theory the necessary research tools that fit better with the research subject (the archaeological material and its interpretation).

6. Quantifying uncertainty of a 3D model

In the following section some mathematical methods are proposed to summarise all the numerical data obtained as reliability and importance indices.

6.1. Assigning reliability and importance to the model's components

As already mentioned, the first step in quantifying the reliability of a 3D model is to build a table (Figure 4) in which all components of the model are detailed, together with their attributes, on the base of the established criteria (see above).

evaluation categories		existence	material	technique	size	texture	position	quantity
WALLS	reliability	1	0,9	0,9	0,8	0,7	1	1
	importance	1	0,9	0,8	0,8	0,3	1	1
FLOORS	reliability	1	1	1	0,8	0,8	0,9	1
	importance	0,6	0,5	0,6	0,8	0,5	0,8	1
ROOF	reliability	0,6	0,7	0,7	0,3	0,7	0,8	1
	importance	0,8	0,7	0,7	0,6	0,7	0,9	1
WINDOWS	reliability	0,7	0,7	0,7	0,3	0,7	0,1	0,3
	importance	0,6	0,4	0,4	0,5	0,4	0,5	0,6
DOORS	reliability	1	0,7	0,7	0,7	0,7	1	1
	importance	0,7	0,6	0,6	0,5	0,6	0,7	0,5
GRINDING STONES	reliability	1	0,7	0,7	0,8	1	0,1	0,9
	importance	0,4	0,3	0,3	0,3	0,3	0,5	0,4
OTHER ARTEFACTS	reliability	0,8	1	1	0,9	0,9	0,4	0,3
	importance	0,1	0,5	0,5	0,5	0,8	0,5	0,5

CONCEPTUAL CATEGORIES		division of space	use of space	social life
	reliability	0,4	0,4	0,2
	importance	0,1	0,1	0,1

Figure 4: *Model's components and subcategories with attributes*

A first view at the relationship between reliability and importance of the components may be achieved by plotting these values on a scattergram (Figure 5). The graphic visualization enables to see the tendency of the components' reliability and importance, enabling a quick visual evaluation of the 3D model's (un)certainty. Note upper left square indicating categories (components of the model) with a high reliability but a low importance, against the lower right corner, important components of the model but with a low reliability. These extremes might be re-evaluated and accordingly treated. The fact that in our case most values are concentrated in the upper right corner (high reliability and high importance) means that there is a consistency in the modeling process and information is generally reliable and open to a critic evaluation.

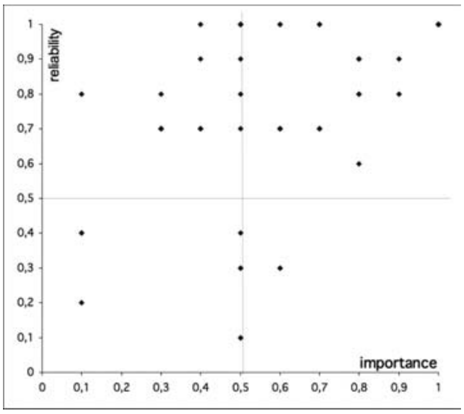


Figure 5: *Scattergram of importance/reliability indices*

6.2. Summarizing the categories

In the next step for each category one general value of reliability and one general value of importance might be computed. As an overall reliability index the average value is proposed because it should consider all the reliabilities of the subcategories. The importance index instead should indicate the detail's most important subcategory (or subcategories), so the maximum function will be used. As the main goal of this paper is to present a solution for evaluating the reliability of a 3D model and to enable its further visual analysis by means of layers defined by given parameters, it is convenient to combine these reliability and importance indices into one variable, so that we obtain one evaluation index for each category.

A required function to compute this had to satisfy two main conditions: a. the values should belong to the $[0,1]$ interval and b. it should distinguish reliability and important indices as variables, so that a cross - change between them would lead to a different result. We propose a following function:

$$R_n = r_n(\sin\frac{\pi}{2}i_n), \tag{1}$$

where r_n is the overall reliability and i_n is the overall importance of the category, computed in the previous paragraph.

A character of the proposed function causes more attention to be paid to the reliability, than the importance index, that is, categories with a high reliability index generally get higher overall value than those of a low reliability even when the latter being of high importance, which is demonstrated on the following graph (Figure 6).

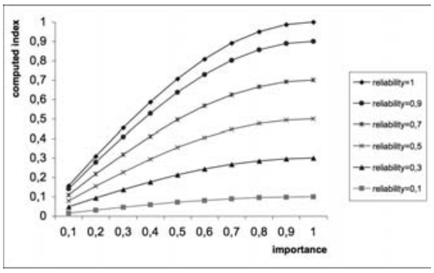


Figure 6: *Example values of the R_n function*

The focus was put on reliability because of the goal of our work, which was to create a reliable model, even when having a lower number of details. Of course exchanging r_n with i_n in the formula would inverse the interpretation. In the following table computed category indices for our model are presented (Figure 7).

Evaluation categories		Mean	Max	R
WALLS	reliability	0,90		0,900
	importance		1,00	
FLOORS	reliability	0,93		0,929
	importance		1,00	
ROOF	reliability	0,69		0,686
	importance		1,00	
WINDOWS	reliability	0,50		0,405
	importance		0,60	
DOORS	reliability	0,83		0,738
	importance		0,70	
GRINDING STONES	reliability	0,74		0,525
	importance		0,50	
OTHER ARTEFACTS	reliability	0,76		0,720
	importance		0,80	
CONCEPTUAL CATEGORIES	reliability	0,33		0,052
	importance		0,10	
average				0,619

Figure 7: *Computed indices of model's components*

6.3. Computing the overall value of the model

Since the above-computed indices already contain all the information about details' reliability and importance, we can simply compute their average and assume it being the value index for the model. So the overall index R was defined as follows:

$$R = \frac{\sum_{n=1}^{N} R_n}{N}, \tag{2}$$

where N is the total number of categories in the model. In our case $R = 0{,}619$, which should be interpreted as a good result for our research objectives. Future evaluations of other models will enable us to define the minimum acceptable value of R of scientifically reliable a 3D model.

6.4. Defining other versions of chosen details

Let's say we want to analyse other versions of the model checking how it goes with different details chosen. Of course these exchanged details may have different parameters - reliability and importance.

To illustrate the following step two hypothetical versions of our model were defined. In the first one all the features were kept as they were before, only excluding conceptual categories, which were very uncertain. In the second we change a bit the concept of the model, defining artefacts and conceptual categories in a more reliable way, but having lower reliability for the walls and windows (separated here into two categories - Window 1 and Window 2). We perform first three steps of the algorithm to all of the newly defined versions obtaining results as follows (Figure 8).

evaluation categories	R (basic model)	R2 (second version)	R3 (third version)
WALLS	0,9	0,9	0,678
FLOORS	0,929	0,929	0,929
ROOF	0,686	0,686	0,686
WINDOW 1	0,405	0,405	0,384
WINDOW 2			0,184
DOORS	0,738	0,738	0,738
GRINDING STONES	0,525	0,525	0,525
OTHER ARTEFACTS	0,720	0,720	0,820
CONCEPTUAL CATEGORIES	0,052		0,184
average	0,619	0,700	0,570

Figure 8: *Reliability of alternative models*

6.5. Final evaluation of the model

Since from the very beginning of the evaluation process fuzzy indices were used as a way of describing the model's details, they can now be used for defining a fuzzy set describing membership of reconstructed categories to the entire model. A formal definition of a fuzzy set M_i (a set of elements belonging to a particular i- version of the model) should be given as follows:

$$M_i = \{(c_n, \quad \mu_i(c_n)), \quad c_n \in M\}, \qquad n = 1,\ldots,N \tag{3}$$

where M is the set of all the categories defined in the model, N is a number of elements of M and $\mu_i : M \to [0,1]$,

$\mu_i(c_n) = R_n^i$ is a membership function (R_n^i is the R_n function given for the i- version of the model). Actually, three fuzzy sets: M_1, M_2, M_3 were already presented in (Figure 8), with elements of M listed in the category column and membership function for all of the sets M_1, M_2, M_3 defined in R, R_2, R_3 columns respectively. The sets can be also represented in a more easy-to-see way by means of 3D bars (Figure 9).

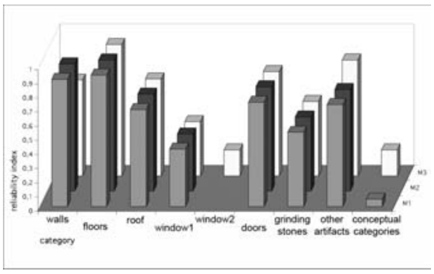


Figure 9: *Visual representation of the fuzzy sets M_1, M_2, M_3*

As one can observe, introduction of fuzzy sets into the process of evaluation of a 3D model provides some new possibilities to express its value by means of numbers and traditionally statistical tools like diagrams and plots. Although it is still only static (but meaningful) information, it is highly desired to improve its use by applying in dynamical visualisation not only of the metadata, but of the model itself.

7. Conclusions

3D visualization is an optimal tool for virtually re-creating and thus analysing archaeological material in its "real" context. It also allows to visually tackle the archaeological remains with the researcher's "mental model", expressed by the (3D) virtually reconstruction of these remains and their context. Accordingly, a 3D model may be viewed as a platform gathering outcomes of a multi-disciplinary research. Thus, a 3D model can be viewed as an expression of the sum up of the acquired knowledge regarding the investigated subject, combining both the perceptions (the archaeological data) with the inferences (the information derived from analysing this data), and thus visually communicating the cognitive process of the archaeological reasoning.

In order to analyse the model from a scientific point of view, aspects of data transparency and reliability of reconstruction should be clearly expressed. The article introduced two key elements that may advance these issues - reliability and importance indices. These indices, obtained by applying concepts deriving from the foundations of fuzzy sets theory, allow a numerical expression of the (un)certainty of a 3D model.

The theoretical discussion presented above will be continued in future work, by developing an algorithm and software that would implement a contingency threshold to a viewer and thus visualize only components passing desired reliability. In the case of Qalqal Asba, the dynamical visualization of the model would not only allow to present its other possible architectural reconstruction (based on known different features from other sources of data), but also give a possibility to show alternaive interpretation of the use of space and organization of the activities in the house.

References

[ACN03] ARNOLD D., CHALMERS A., NICCOLUCCI F.: VAST 2003. *Eurographics Association, Aire-la-Ville, Switzerland* (2003).

[All98] ALLEN W. J.: Levels of archaeological evidence illustrated in three-dimensional computer models. *Master Thesis, Mississipi State University* (1998).

[BDF*02] BARD K. A., DIBLASI M. C., FATTOVICH R., MANZO A., PERLINGIERI C., PHILLIPSON L., SULAS F.: The IOU and BU Archaeological Expedition at Bieta Giyorghis (Aksum): 2002 Field season. *Nyame Akuma*, 57 (2002), 2–6.

[BFS00] BARCEL J. A., FORTE M., SANDERS D. H.: Virtual reality in archaeology. *Oxford, BAR International Series 843* (2000).

[BP04] BEEX W., PETERSON J.: The Arminghal henge in space and time: how virtual reality contributes to research on its orientation. *Enter the Past - the E-Way into the Four Dimensions of Cultural Heritage Proc. CAA2003 Oxford, BAR International Series 1127* (2004), 490–3.

[Bur02] BURENHULT G.: Archaeological Informatics: Pushing the Envelope Proc. CAA 2001. *Oxford, BAR International Series 1016* (2002).

[CCSN04] CHRYSANTHOU Y., CAIN K., SILBERMAN N., NICCOLUCCI F.: VAST 2004. *Eurographics Association, Aire-la-Ville, Switzerland* (2004).

[Cho03] CHOI D.: Enhancing the power of Web search engines by means of fuzzy query. *Decision Support Systems 35(1)* (2003), 31–44.

[Dol94] DOLAN C.: The magic of fuzzy logic. *Global Investor 74:11* (1994).

[DS03] DOERR M., SARRIS A.: The Digital Heritage of Archaeology, Proc. CAA 2002. *Archive of Monuments and Publications, Hellenic Ministry of Culture* (2003).

[EFH76] EKSTROM R. B., FRENCH J. W., HARMAN H. H.: *Manual of Kit of Factor-Referenced Cognitive Tests*. Princeton, Educational Testing Service, 1976.

[FAF*00] FRISCHER B., ABERNATHY D., FAVRO D., LIVERANI P., DE BLAAUW S.: Virtual Reality and Ancient Rome: The UCLA Cultural VR Labs Santa Maria Maggiore Project. *Virtual Reality in Archaeology Oxford, BAR International Series 843* (2000), 155–62.

[Fis00] FISHER P.: Sorites paradox and vague geographies. *Fuzzy Sets and Systems 113(1)* (2000), 7–18.

[FNRB02] FRISCHER B., NICCOLUCCI F., RYAN N. S., BARCEL J. A.: From CVR to CVRO: the past, present, and future of cultural virtual reality. *Virtual Archaeolog Proc. VAST Arezzo, Oxford, BAR International Series 1075* (2002), 7–18.

[For00] FORTE M.: About virtual archaeology: disorders, cognitive interactions and virtuality. *Virtual Reality in Archaeology Oxford, BAR International Series 843* (2000), 247–59.

[FS97] FORTE M., SILIOTTI A.: *Virtual Archaeology - Great Discoveries Brought to Life Through Virtual Reality*. London, Thames & Hudson, 1997.

[Hal97] HALL A. V.: A generalized taxon concept. *Botanical Journal of the Linnean Society 125* (1997), 169–180.

[HN03] HERMON S., NICCOLUCCI F.: A fuzzy logic approach to typology in archaeological research. *The Digital Heritage of Archaeology Athens, Archive of Monuments and Publications* (2003), 307–310.

[HND05] HERMON S., NICCOLUCCI F., D'ANDREA A.: Some evaluations on the potential impact of virtual reality on the archaeological scientific research. *Proc. VSMM 2005, Ghent, Belgium* (2005), 105–14.

[HSB93] HARRIS T. R., STODDARD S. W., BEZDEK J. C.: Application of fuzzy set clustering for regional typologies. *Growth and Change 24(2)* (1993), 155–165.

[Kan00] KANTER J.: Realism vs. reality: creating virtual reconstructions of prehistoric architecture. *Virtual Reality in Archaeology* (2000), 47–52. BAR International Series, 843.

[Kos93] KOSKO B.: *Fuzzy Thinking: The New Science of Fuzzy Logic*. Hyperion, 1993.

[MF93] MCNEILLL D., FREIBERGER P.: *Fuzzy Logic*. Simone and Shuster, 1993.

[Mor93] MORACZEWSKI I. R.: Fuzzy logic for phytosociology II. Generalizations and predictions. *Vegetatio 106*, 1 (1993), 13–20.

[MPPF03] MANZO A., PERLINGIERI C., PERLINGIERI R., FOCCILLO M.: Real space beyond solid models: Spatial metadata in ethnoarchaeology. *The Reconstruction of Ancient Landscapes through Digital Technologies* (2003), 71–80. (BAR International Series, 1151).

[Muc97] MUCINA L.: Classification of vegetation: past, present and future. *Journal of Vegetation Science 8*, 6 (1997), 751–760.

[NH06] NICCOLUCCI F., HERMON S.: Beyond the artifact. *Proc. CAA2004* (2006). (Archaeolingua).

[Nic02] NICCOLUCCI F.: Virtual archaeology. *Proc. VAST, Arezzo* (2002). (BAR International Series, 1075).

[Nov89] NOVAK V.: *Fuzzy Sets and their Applications*. Adam Hilger, 1989.

[OLGE98] OLANO M., LOIDI J. J., GONZALES A., ESCUDERO A.: Improving the interpretation of fuzzy partitions in vegetation science with constrained ordinations. *Plant Ecology*, 134 (1998), 113–118.

[PMP] PERLINGIERI C., MANZO A., PERLINGIERI R.: Archaeological data, past living spaces, and present habits: the virtuality possible. *Digital Discovery and Exploring New Frontiers in Human Heritage, Proc. CAA2006, Fargo*. (in press).

[Rag00] RAGIN C. C.: *Fuzzy-Set Social Science*. University of Chicago Press, 2000.

[Rap93] RAPOSA M.: The fuzzy logic of religious discourse. *American Journal of Semiotics*, 10 (1993), 1–101.

[Rei89] REILLY P.: Data visualization in archaeology. *IBM Systems Journal 28*, 4 (1989), 569–579.

[Rob86] ROBERTS D. W.: Ordinations on the basis of fuzzy set theory. *Vegetatio 66* (1986), 123–134.

[Rob89] ROBERTS D. W.: Fuzzy systems vegetation theory. *Vegetatio 83* (1989), 71–80.

[RR97] ROBERTS J. C., RYAN N. S.: Alternative archaeological representations within virtual worlds. Fourth VR-SIG 97 Conference.

[RS89] REILLY P., SHENNAN S.: Applying solid modeling and animated three-dimensional graphics to archaeological problems. *Technical Report UKSC 209* (1989).

[Rui99] RUIZ B.: Science and philosophy. *Diogenes 188* (1999), 73–84.

[Rya96] RYAN N. S.: Computer-based visualisation of the past: technical realism and historical credibility. *Imaging the Past: Electronic Imaging and Computer Graphics in Museums and Archaeology*, 114 (1996), 95–108. (British Museum Occasional Papers).

[Sat96] SATTLER R.: Classical morphology and continuum morphology: opposition and continuum. *Annals of Botany 78* (1996), 577–581.

[Sim97] SIMS D.: Archaeological models: pretty pictures or research tools? *IEEE Computer Graphics 17*, 1 (1997), 13–15.

[Vat02] VATANEN I.: Argumentation paths in information infrastructure of the archaeological virtual realities. *Archaeological Informatics: Pushing the Envelope, Proc. CAA 2001* (2002), 494–99. (BAR International Series, 1016).

[Zad65] ZADEH L. A.: Fuzzy sets. *Information and Control 8* (1965), 338–353.

[Zim84] ZIMMERMAN H. J.: *Fuzzy Set Theory and its Applications*. Kluwer-Nijhoff Publishing, 1984.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

A Concept for the Separation of Foreground/Background in Arabic Historical Manuscripts using Hybrid Methods

W. Boussellaa[1] and H. El Abed[2] and A. Zahour[3]

[1]REsearch Group on Intelligents Machines (REGIM), ENIS, University of Sfax, Tunisia
[3]Institut for Communications Technology (IfN), Braunschweig Technical University, Germany
[2]Equipe Gestion Électronique de Document (GED), University of Le Havre, France

Abstract

This paper presents a new color document image segmentation system suitable for historical Arabic manuscripts. Our system is composed of a hybrid method which couple together background light intensity normalization algorithm and k-means clustering with maximum likelihood (ML) estimation, for foreground/ background separation. Firstly, the background normalization algorithm performs separation between foreground and background. This foreground is used in later steps. Secondly, our algorithm proceeds on luminance and distort the contrast. These distortions are corrected with a gamma correction and contrast adjustment. Finally, the new enhanced foreground image is segmented to foreground/background on the basis of ML estimation. The initial parameters for the ML method are estimated by k-means clustering algorithm. The segmented image is used to produce a final restored document image.

The techniques are tested on a set of Arabic historical manuscripts documents from the National Tunisian Library. The performance of the algorithm is demonstrated on by real color manuscripts distorted with show-through effects, uneven background color and localized spot.

Categories and Subject Descriptors (according to ACM CCS): I.7.5 [Document and Text Processing]: Document Capture/ Document analysis

1. Introduction

The historical documents, preserved at the National Library of Tunisia, are considered as an important part of Arabic cultural heritage. These funds suffer from a progressive degradation and therefore risk to disappear. The automatic processing of this type of documents in order to restore and use, is a definite advantage which is confronted with many difficulties due to the storage condition and the complexity of their content. In fact, historical documents have many particularities which hinder classical color document image segmentation algorithms. Figure 1 illustrates the most common deteriorations that appeared in historical Arabic document images which are: The show-through effects (Figure 1 left), the presence of spot due to the humidity absorbed by paper, and an uneven background color paper (Figure 1 right), the presence of fold and tear, and the distortions due to the natural curvature of pages.

Most previous document image enhancement algorithms have been designed primarily for binarization of modern documents. These methods aim to extract text from noisy documents with uneven background. Three popular methods, namely Otsu's thresholding technique [Ots79], entropy techniques proposed by Kapur and al. [KSW85] and the minimal error technique by Kittler and Illingworth [KI86], are analysed and compared in [LYT*03, LVPG02]. Another entropy-based method specially designed for historical document segmentation [CR00] deals with the noise inherent in the paper especially in documents written on both sides. Wang and al. [WXLT03] presented methods to separate text from background noise and bleed-through text (from the backside of the paper) using direct image matching and directional wavelets. Other methods for historical document image enhancement are driven by the goal of improving human readability of the documents [CR02].

This paper presents a new method for foreground/background segmentation of color historical

Arabic manuscripts. This method combines two techniques of segmentation: The foreground/background separation with background light intensity normalization algorithm and the improvement of the obtained result on the basis of ML estimation after contrast adjustment with gamma correction and histogram normalization of the foreground image. The following paper describes the proposed method and experimental results.

Figure 1: *Arabic historical documents image: (left) Show-through effects, (right) uneven background*

2. State of the Art

The works described in [BHH*98, GPH05, LBE04] are dedicated for foreground/background separation of color document images. DjVu [BHH*98] implements an efficient foreground/background separation in the context of compression. The approach is based on a multi-scale bi-color clustering that considers several grids of increasing resolution. The technique works well for a large class of documents (gray as well as color) but fails for documents with low contrast.

Leydier and al [LBE04] have achieved an adaptative algorithm for the segmentation of color images suited for document image analysis. The algorithm is based on a serialization of the k-means algorithm that is applied sequentially by using a sliding window over the image. The algorithm reuses information about the clusters computed by the previous classification and automatically adjusts the clusters during the windows displacement in order to better adapt the classifier to any new local modification of the colors. The used colorspaces are RGB and HSL (Hue, Saturation, and Luminosity).

Garain and al [GPH05] have proposed an adaptive method for foreground/background separation in low quality color document images. A connected component labelling is initially implemented to capture the spatially connected similar color pixels. Next, Dominant background components are determined to divide the entire image into a number of grids each representing local uniformity in illumination background. Finally, foreground parts are located using local information around them. This method achieved good results compared to DjVu [BHH*98].

Shi and al [SG04, SG05] have proposed a color document image enhancement algorithm of palm leaf manuscripts. This method is based on background light intensity normalization. The background approximation is designed to overcome the unevenness of the document background and the low contrast. The techniques are tested on a set of palm leaf images from various sources and the results show significant improvement in readability.

3. Proposed Methodology

The proposed document enhancement methodology permits the improvement of the quality of historical Arabic manuscripts which presented uneven background and low contrast due to the traditional mode of manufacture and the effect of ageing and degradation. It consists of the following steps: foreground extraction, contrast adjustment, foreground/background segmentation, reconstruction of document image with smoothing. The developed document segmentation method operates with background light intensity normalization algorithm proposed by shi and al [SG04, SG05] and applied to palm leaf manuscripts. We have improved this technique with the histogram normalization used in color image manuscript context. The segmentation method proceeds on luminance and distort the contrast. These distortions are corrected with a gamma correction and contrast adjustment. The new enhanced foreground image is segmented to foreground/background on the basis of ML estimation. The initial parameters for the ML method are estimated by k-means clustering algorithm. The segmented image is used to produce a final restored document image. Figure 2 presents the flowchart of our proposed methodology.

The steps below are described in the following sections:

- Application of an iterative background light intensity normalization algorithm for a first foreground/background separation.
- Correction of visual distortions of obtained foreground using gamma correction and histogram normalization.
- Estimation of the parameters with K-means algorithm for ML method. This algorithm performs final foreground/background segmentation.
- Reconstruction of images color space and production of the restored manuscript.

3.1. Background Light Intensity Normalisation Algorithm

Background light intensity normalization algorithm is applied on historical manuscripts documents presented an uneven background and low contrast. Therefore, the choice of color space is important. This technique performs background approximation at first. Secondly, foreground normalisation is carried out from approximated background Y_{Back} and luminance image $Y_{Original}$ of the original image. Figure 3 shows the normalization process.

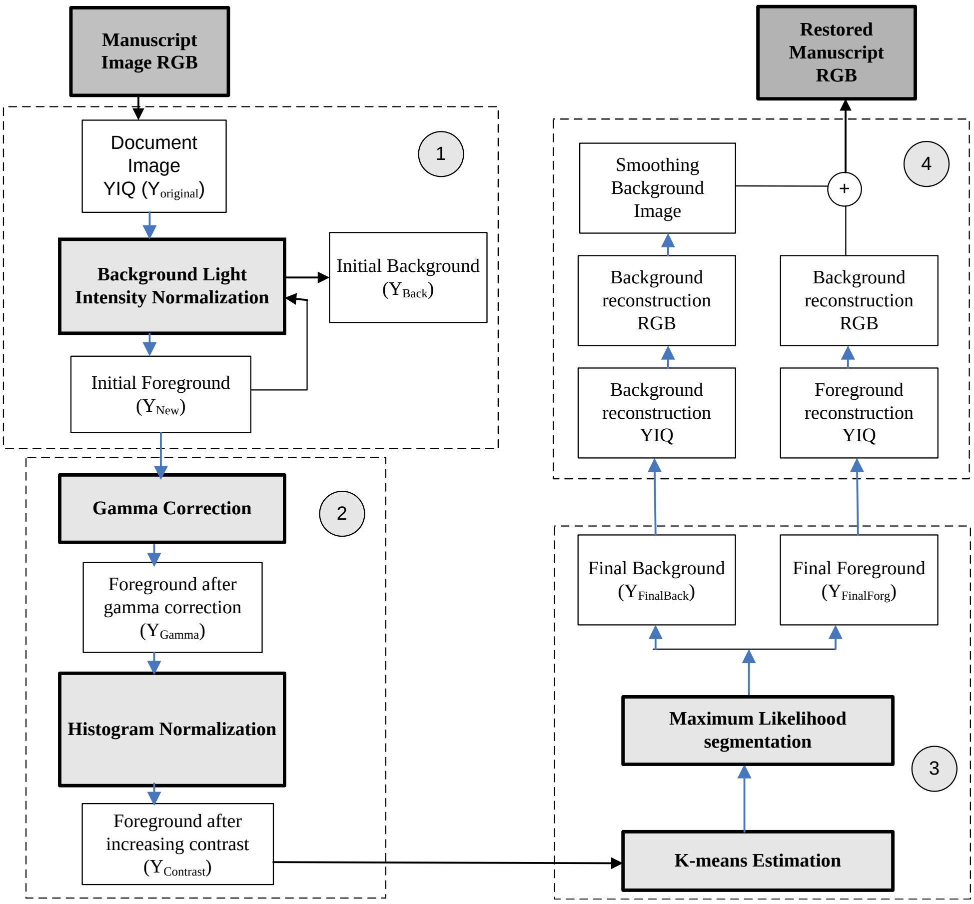


Figure 2: *Flowchart of proposed methodology*

3.1.1. Feature Choices

The choice of YIQ (Y: luminance channel; I and Q: chrominance color channels) colorspace is justified by the fact that the human vision is very sensitive to the change of luminosity. Moreover, the variation in light intensity caused by the uneven background of historical manuscripts is captured in L channel. An example of image decomposition from RGB to YIQ colorspace is presented in Figure 4.

3.1.2. Background Approximation

Background approximation algorithm start with a global binarization of the Y channel using Otsu method [Ots79]. This technique computes a global threshold for text extraction based on minimizing the intraclasses variance of the image's pixels. The steps of the background approximation algorithm are presented below:

- Computing the horizontal projection profile H of binary document image from L channel.
- Computing the average of histogram values M (step1).
- Scanning the image line by line $Y_{Original}$ and background approximation Y_{Back} (step 2).
- Recursive estimation of each final pixel grayscale of the image Y_{Back} (step3).

After experimental work, for the case of historical manuscripts, we suggest the following parameter values: $window\ size = 3 \times 3$, $mtime = 20$. An example of a resulting background approximation is given in figure 4.

3.1.3. Image Normalisation

Foreground normalisation is obtained from L channel of original image and estimated background Y_{Back}. The light intensity pixels values of the new foreground Y_{New} are computed according to the following formulas, equation 1 and equation 2:

- Linear normalisation by translation

$$Y_{New} = (Y_{Original} - Y_{Back}) + C \tag{1}$$

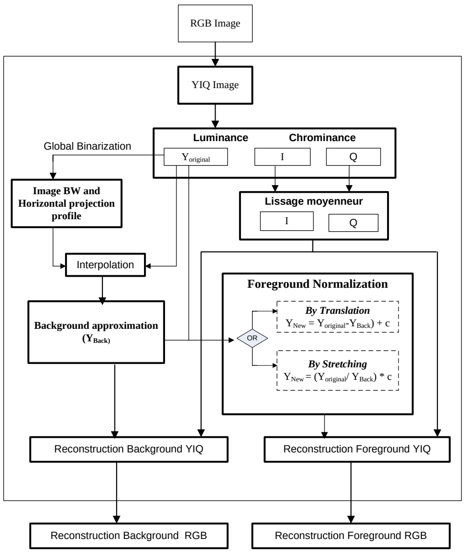


Figure 3: *Background light intensity normalisation process*

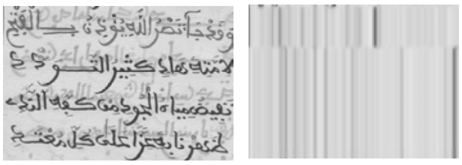


Figure 4: *Background light intensity normalisation on luminance channe Y: (left) Manuscript image, (right) Approximated background*

- Linear normalisation by stretching

$$Y_{New} = (\frac{Y_{Original}}{Y_{Back}}) \cdot C \tag{2}$$

To ensure that the value does not exceed 255, C is set to 255 (usually) to make the background color white. The resultant normalized images are shown in Figure 5.

The normalisation by stretching is more adapted for the case of manuscript documents which present an uneven background and a low contrast. Experimental works show that iterating the normalisation process is necessary. We suggest three iterations which gives sufficient results. The obtained foreground Y_{New} is used in the later steps.

3.2. Histogram Transformations

3.2.1. Gamma Correction

Cheng and al, and Tremeau and al [CJSW01,TFMB04] have shown in their survey on color space, that the image processing with YIQ color space requires a gamma correction. In our case, the coefficient of gamma correction is the ratio between the average of intensity values of original image Yoriginal and the resulting foreground Y_{New}, according to the following formula, equation 3.

Figure 5: *Foreground normalisation: (top) $Y_{Original}$ image, (left) By translation, (right) By stretching*

$$\gamma = \frac{Mean(Y_{New})}{Mean(Y_{Original})} \tag{3}$$

We notice that the values of gamma are usually greater than 1. This operation increases the contrast of image L_{New}. Figure 6 shows the effect of gamma correction.

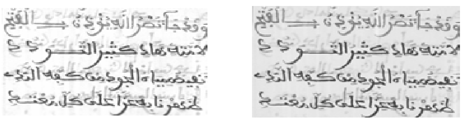


Figure 6: *Gamma correction: (left) Foreground before gamma correction, (right) Foreground after gamma correction*

3.2.2. Histogram Normalisation

After gamma correction, the resulting foreground Y_{Gamma} contains again pale colors. In order to increase the contrast of image, we apply a stretching to the intensity values of image histogram using a proportion value. Then, the image $Y_{Contrast}$ is produced with a proportion between 2% and 8% which gives correct results shown in figure 7.

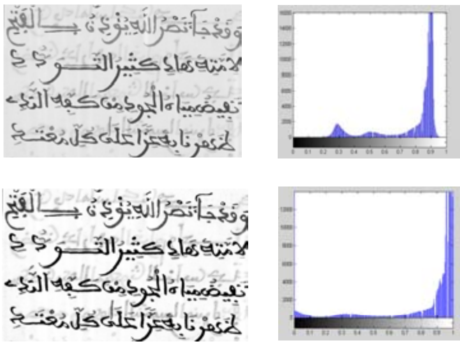


Figure 7: *Histogram normalisation: (top) Foreground after gamma correction,with the intensity histogram, (bottom) Foreground after contrast adjustment, with the intensity histogram*

3.3. Foreground/Background Segmentation

Document image manuscript segmentation can be considered as a statistical classification problem. The estimation of parameters of classification is given by Kmeans algorithm improved by ML method.

3.3.1. Initialisation of K-means Algorithm

K-means algorithm operates on the image $Y_{Contrast}$. This technique computes the statistical features vectors for the foreground/background classification. K-means algorithm performs a first foreground/background classification. Figure 8 illustrates the result of segmentation.

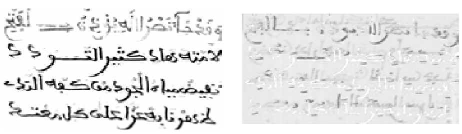


Figure 8: *K-means foreground/background segmentation: (left) Foreground, (right) Background*

We notice that there is a significative loss of text information from foreground. In order to improve results of final segmentation, we refine the parameters of classification estimated by k-means algorithm by using ML algorithm.

3.3.2. Maximum Likelihood Algorithm

The maximum likelihood classifier is one of the most popular methods of classification in remote sensing, in which a pixel with the maximum likelihood is classified into the corresponding class. The likelihood L_k is defined as the posterior probability of a pixel belonging to class K, and it computed as the following, equation 4 .

$$L_k = P(K|X) = \frac{P(K) \cdot P(X|K)}{\sum P(i) \cdot P(X|i)} \tag{4}$$

Where:

- $P(K)$: prior probability of class k
- $P(X|K)$: conditional probability to observe X from class k, or probability density function

Therefore L_k depends on $P(X|k)$ or the probability density function. Foreground/background segmentation of Ycontrast image is performed by ML method. This technique relies on the likelihood function of the distribution of image intensity pixels. The ML method estimates the probability that a pixel belongs to its corresponding class which is foreground or background and assigns it when its probability is maximal. We are using two probability distributions, the Gaussian law and the Raleigh law. According to the distribution, the likelihood L_k is expressed in the following equations 5 and 6.

- L_k according to Gaussian distribution:

$$L_{k=1,2}(Y_{Contrast}) = \frac{1}{\sigma_k \sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma_k^2} \cdot (Y_{Contrast} - \mu_k)^2} \tag{5}$$

- L_k according to Raleigh distribution:

$$L_{k=1,2}(Y_{Contrast}) = \frac{1}{\mu_k \sqrt{\frac{2}{\pi}}} \cdot e^{-\frac{Y_{Contrast}^2}{2(\mu_k \sqrt{\frac{2}{\pi}})^2}} \tag{6}$$

The pixel j in $Y_{Contrast}$ is labelled L_{k_j} according to the following equation 7.

$$L_{k_j} = \max_x (L_k(Y)) \tag{7}$$

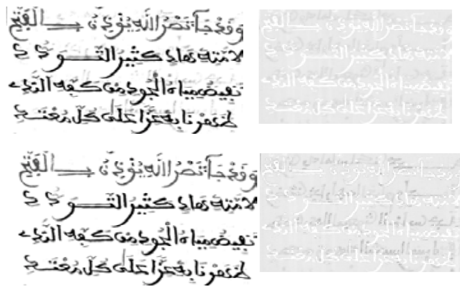


Figure 9: *ML Foreground background segmentation: (left) Gaussian distribution, (right) Raleigh distribution*

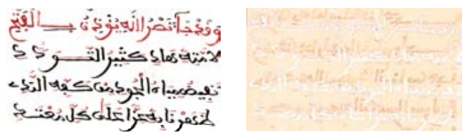


Figure 10: *Foreground/background RGB Reconstruction: (left) Foreground, (right) Background*

Experimental work shows that for the case of historical manuscripts, the Raleigh distribution gives better results for foreground/background segmentation. Figures 9 and 10 illustrate results in HSL and RGB color spaces.

3.4. Manuscripts Restoration

Foreground/background segmentation by ML method is used for the restoration of historical manuscripts. In fact, the restored image is constructed by superposition of the foreground and the average of background in RGB colorspace. Figure 11 illustrates visually the restored historical manuscript.

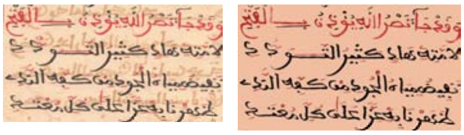


Figure 11: *Manuscript output: (left) Original, (right) Restored*

4. Conclusion and Future Works

In this paper, we have presented a hybrid method for foreground/background segmentation suited for Arabic documents manuscripts distorted with show-through effects and uneven background. This technique is based on four steps: (a) foreground extraction with iterative light intensity normalisation algorithm, (b) postprocessing of obtained foreground with double contrast adjustment, (c) Foreground/background segmentation with ML method, (d) reconstruction of restored manuscript document.

Our future objective aims to perfect our method. We are planning to assist the user to define automatically the iterations number of the normalisation process. Moreover, we aim to improve the texture segmentation method in order to classify the document into three types of texture information: Text, background, and graphic, can be useful to an indexing and retrieval system of Arabic historical documents.

Acknowledgment

The Authors will Thanks Prof. Adel Alimi, Director of the Ecole Nationale des Ingénieurs de Sfax - ENIS, Head of the Laboratory REGIM. Special thanks to Prof. Abdellattif Benabdehafid, director of the GED (Equipe Gestion Électronique de Document), University of Le Havre. Special thanks to Dr. Volker M"argner, Head of the image processing group, at the Department of Signal Processing for Mobile Informations Systems , Institute for Communications Technology (IfN), Braunschweig Technical University.

References

[BHH*98] BOTTOU L., HAFFNER P., HOWARD P., SIMARD P., BENGIO Y., LECUN Y.: High quality document image compression with djvu. *Journal of Electronic Imaging 7*, 3 (July 1998), 410–425.

[CJSW01] CHENG H., JIANG X., SUN Y., WANG J.: Color image segmentation:advances and prospects. In *Pattern Recognition* (2001), vol. 34, pp. 2259–2281.

[CR00] C.A.B.MELLO, R.D.LINS: Image segmentation of historical documents. *Visual2000* (September 2000).

[CR02] C.A.B.MELLO, R.D.LINS: Generation of images of historical documents by composition. *ACM Symposium on Document Engineering* (2002).

[GPH05] GARAIN U., PAQUET T., HEUTTE L.: On foreground-background separation in low quality color document images. In *Eighth International Conference on Document Analysis and Recognition (ICDAR)* (2005), pp. 585–589.

[KI86] KITTLER J., ILLINGWORTH J.: Minimum error thresholding. *Pattern Recognition 19*, 1 (1986), 41–47.

[KSW85] KAPUR J. N., SAHOO P. K., WONG A. K. C.: A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing 29* (1985), 273–285.

[LBE04] LEYDIER Y., BOURGEOIS F. L., EMPTOZ H.: Serialized k-means for adaptative color image segmentation: Application to document images and others. In *Document Analysis Systems (DAS)* (2004), pp. 252–263.

[LVPG02] LEEDHAM G., VARMA S., PATANKAR A., GOVINDARAYU V.: Separating text and background in degraded document images - a comparison of global threshholding techniques for multi-stage threshholding. In *Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02)* (2002), p. 244.

[LYT*03] LEEDHAM G., YAN C., TAKRU K., TAN J. H. N., MIAN L.: Comparison of some thresholding algorithms for text/background segmentation in difficult document images. In *Seventh International Conference on Document Analysis and Recognition (ICDAR)* (2003), pp. 859–869.

[Ots79] OTSU N.: A threshold selection method from gray level histogram. *IEEE Transactions in Systems, Man, and Cybernetics 9* (1979), 62–66.

[SG04] SHI Z., GOVINDARAJU V.: Historical document image enhancement using background light intensity normalization. *The 17th International Conference on Pattern Recognition (ICPR) 01* (2004), 473–476.

[SG05] SHI Z., GOVINDARAJU V.: Historical document image segmentation using background light intensity normalization. *SPIE Electronic Imaging, Document Recognition and Retrieval XII 5676*, 1 (2005), 167–174.

[TFMB04] TRÉMEAU A., FERNANDEZ-MALOIGNE C., BONTON P.: *Image numérique couleur de l'acquisition au traitement*. Dunod, Paris, 2004.

[WXLT03] WANG Q., XIA T., LI L., TAN C. L.: Document image enhancement using directional wavelet. *Computer Vision and Pattern Recognition (CVPR) 02* (2003), 534.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Procedural 3D Reconstruction of Puuc Buildings in Xkipché

Pascal Müller [1] Tijl Vereenooghe [2] Peter Wonka [3] Iken Paap [4] Luc Van Gool [1,2,5]

[1]Computer Vision Lab, ETH Zürich, Switzerland
[2]Center for Archaeological Sciences / ESAT, K.U. Leuven, Belgium
[3]PRISM - Partnership for Research in Spatial Modeling, Arizona State University, USA
[4]Institut für Griechische und Lateinische Philologie, Romanistik und Altamerikanistik, University of Bonn, Germany
[5]PSI / ESAT, K.U. Leuven, Belgium

Abstract

This paper examines how architectural shape grammars can be used to procedurally generate 3D reconstructions of an archaeological site. The Puuc-style buildings found in Xkipché, Mexico, were used as a test-case. We first introduce the ancient Mayan site of Xkipché and give an overview of the building types as distinguished by the archaeologists, based on excavations and surveys of the building remains at the surface. Secondly, we outline the elements of the building design that are characteristic of the Puuc architecture. For the creation of the actual building geometries, we further determine the shape grammar rules for the different architectural parts. The modeling system can then be used to reconstruct the whole site based on various GIS (Geographical Information Systems) data given as input, such as building footprints, architectural information, and elevation. The results demonstrate that our modeling system is, in contrast to traditional 3D modeling, able to efficiently construct a large number of high quality geometric models at low cost.

Categories and Subject Descriptors (according to ACM CCS): F.4.2 [Mathematical Logic and Formal Languages]: Grammars and Other Rewriting Systems I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling I.6.3 [Simulation and Modeling]: Applications J.6 [Computer-Aided Engineering]: Computer-Aided Design (CAD)

1. Introduction

This paper addresses the 3D reconstruction of archaeological sites using procedural modeling. This approach is particularly suited for archaeological purposes, as a historically accurate reconstruction often depends on fragmentary remains and formal architectural "rules" as derived from similar buildings at other sites. In this paper we take Puuc-style architecture at Xkipché in Mexico as a test case. Puuc is a style of Pre-Columbian Maya architecture and tends to be rule-based. Xkipché has been the subject of archaeological research over a period of more than 15 years. This resulted in the detailed mapping of the building remains above ground and those recovered through excavations in several areas of the city.

Being able to reconstruct large cityscapes is an important issue for archaeology. These reconstructions should not be limited to a few major monuments, but also include other buildings, such as domestic quarters. In this context, traditional 3D modeling tools often require too much manual work and their application is therefore overly expensive for archaeological projects. In contrast, our approach proves to be efficient and fairly simple. Furthermore, our procedural modeling approach allows for the testing of several hypotheses by adjusting some of the parameters. This results in a powerful platform for archaeological discussion and exploration. Our contribution is the detailed formal description and grammatical encoding of Puuc stone buildings in Xkipché. The implementation is based on our *CGA Shape* grammar, recently introduced in [MWH*06].

The paper is structured as follows: In section 2 we describe the archaeological background and related work in procedural modeling. Section 3 explains the general formal design of Puuc buildings in Xkipché. In section 4, we give a short introduction to CGA Shape and present the rules to generate the 3D models. The results we achieved are shown in section 5 and discussed in section 6. Our conclusions can be found in section 7.

2. Background and Related Work

Xkipché is located in the Puuc region, a vaguely defined geographical area which includes the southernmost parts of the state of Yucatan as well as an adjoining section of the northern part of the state of Campeche in Mexico. In this region the Puuc architecture attained its ultimate refinement. The ruins of Xkipché are situated about 9.5 kilometers southwest of Uxmal and the city was inhabited from the 6th until the 10th century AD. The area beyond Uxmal was archaeologically not well-known until the late 1980s. Since 1991 researchers of the University Bonn (Institut für Altamerikanistik und Ethnologie) have been carrying out archaeological excavations and a detailed study of the exposed building remains [Pre99]. They manually created a 3D reconstruction of the so-called palace building using CAD software. But these models were not very detailed and none of the surrounding buildings have been created.

Puuc is a subtype of Mayan architecture which is characterized by its veneer-over-concrete construction technique resulting in geometric and repetitive façade structures. Well preserved examples of Puuc-style buildings can be found at Uxmal or Kiuic (see figure 1). In contrast to classical Greek and Roman architecture, the formalization of Puuc or Mayan architecture has received little attention from scholars studying the architectural remains. As a result, the current research could not build upon a formal definition of architectural rules and guidelines (with the exception of some short paragraphs about architecture in [And75, Pol80]). However, comparing Mayan and classical architecture, Andrews explicitly states in [And75] that "The smallest temple, including the platform which supports it, can be observed to be composed of a number of discrete parts, or elements, whose ordering appears to be dependent on a set of rules as explicit as those governing the Roman and Greek orders." This quote supports the feasibility of reconstructing Mayan cityscapes by means of procedural modeling.

Figure 1: *Photograph of a well preserved Puuc-style building in Kiuic, Mexico.*

Procedural modeling techniques for urban environments make use of L-systems [PM01, PL91], shape grammars [Sti75, Sti80, WWSR03, MWH*06], and stack-based languages [Hav05]. In many cases, the procedural modeling techniques are adapted to specific problems, such as modeling of Siza's mass housing [Dua02], medieval houses [BBJ*01], castles [GBHF05], and Chinese temples [LWH*05]. Our work is based on the *CityEngine* software suite [PM01, MWH*06], a comprehensive modeling system that allows to design mass models, create architectural details, and place vegetation. In [MVUG05], the grammar-based reconstruction approach using the CityEngine has been successfully introduced to the archaeological community. Therein, the general CityEngine pipeline for a stochastical creation of buildings has been described, but no detail has been given concerning the architectural rules (for an accurate reconstruction of selected buildings).

3. Puuc Architecture in Xkipché

3.1. Building Types

In Xkipché the archaeological research has identified eight main concentrations of buildings (see groups A-H in figure 2), of which group B probably represents the ceremonial center, whereas groups A, C and D have a residential character. The largest and most important building of the site is located in the center of group A; it is referred to as palace building.

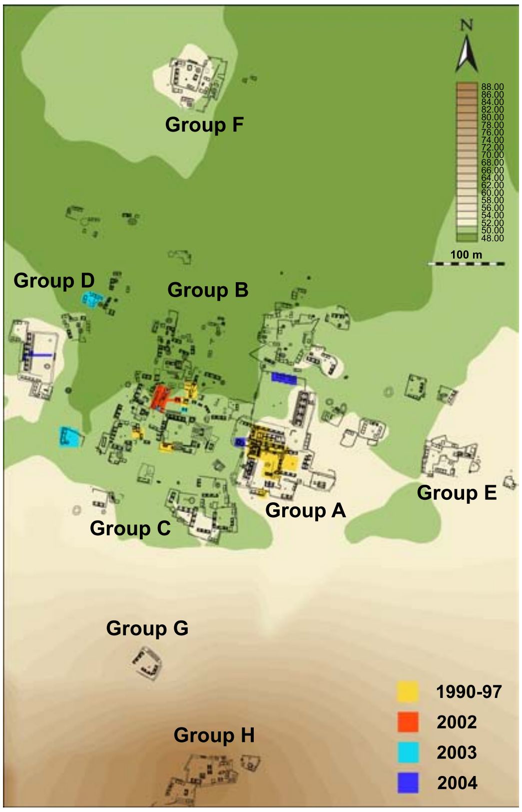


Figure 2: *Overview map of Xkipché with building footprints and elevation. Additionally, the different excavation campaigns over the years are illustrated.*

The archaeologists recognized 18 different building types (illustrated in figure 3). Several of these types, however, only represent one or a part of a building. One of the advantages of Xkipché as a test-case is the level of preservation of the buildings: a lot of the ancient stone buildings are still (partially) standing. Moreover, the shapes of the stones from the respective elements are often very specific. In most cases the archaeologists can easily attribute the architectural fragments, retrieved near collapsed buildings, to their original place in the façade. However, details of the decoration - like e.g. the number and arrangement of colonettes - remain uncertain in many cases. Although the number of stones found can contribute to a more precise reconstruction of the original buildings, one can not neglect the fact that decorated stones of older buildings have frequently been taken away and were reused by the inhabitants themselves.

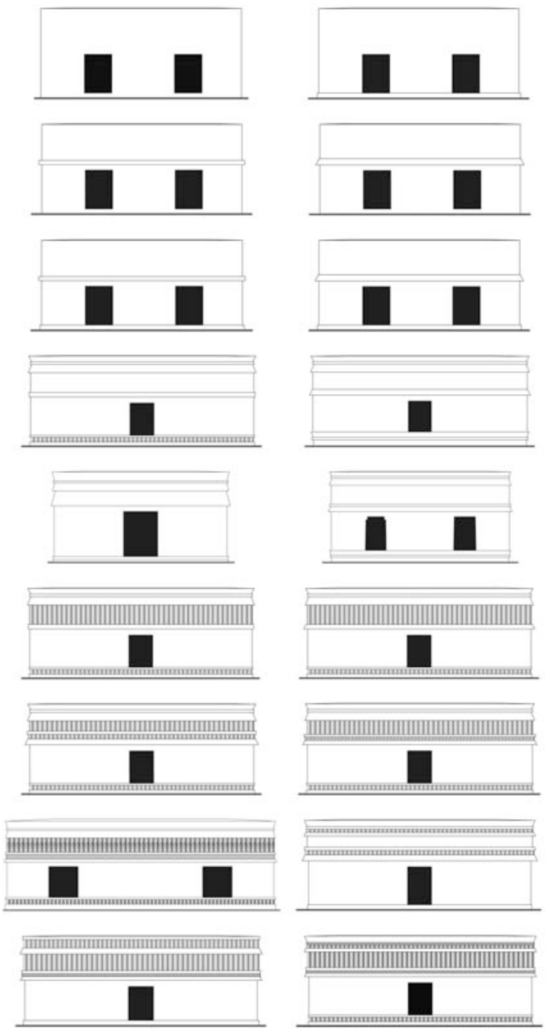


Figure 3: *The 18 building types found in Xkipché. These appearances have been archaeologically reconstructed based on preserved building parts and the remains on the surface.*

3.2. Building Design

Puuc houses were often built on a platform sub-structure, a cut and stucco stone exterior filled with densely packed gravel. The buildings have rubble-filled concrete walls faced by a thin veneer of dressed stone. The dominant characteristics of Puuc-style architecture in Xkipché are a plain lower wall (with openings) above a rather elaborate base molding, and on the upper part of the façade a large medial molding, a frieze (with or without decoration) and a usually high cornice molding [And75, Pol80]. This general building design is illustrated in figure 4 (top).

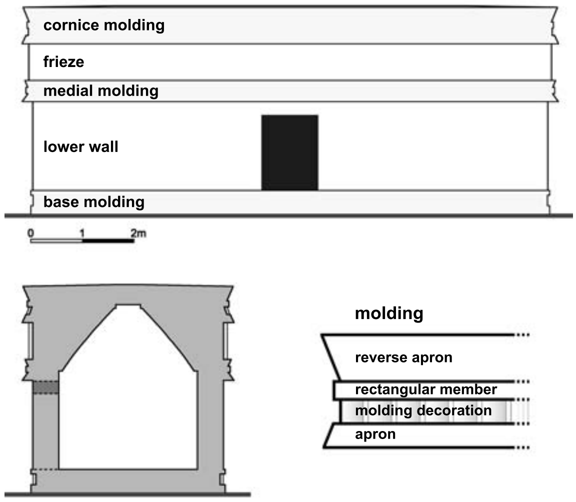


Figure 4: *Top: formal design of a generic stone house in Xkipché. Bottom left: the same building in profile. Bottom right: close up of a four-member molding.*

Apart from the doors, the only wall openings were small rectangular ventilators, often just below the medial molding. Wood was only exceptionally employed as door lintels or unstructurally in the corbelled masonry vaulting system. The average width of the door openings is about 100-120 cm. The door jambs had a strong tendency to incline inwards slightly [Pol80]. Puuc doors were sometimes framed by columns with simple, rectangular capitals [Car86], complemented by small corbels at the top of each jamb.

The frieze decoration (if any) consists usually of colonettes (serrated cylinders) or mosaic elements (lattice work, T-shapes, stylized serpent heads, stepped frets...) of limestone masonry, creating geometric repetition and symmetry. Other typical elements of Puuc façades were long-nosed masks, often supposed to represent the rain god Chac. This more elaborate decoration was frequently found over doorways and at the corner of buildings.

In [Pol80], Pollock gives an overview of moldings, which can have several appearances. Three-member moldings seem to be most common in Puuc architecture, usually consisting of (1) an apron member, a middle rectangular member and an upper reverse apron or (2) an apron member, a decorated member and a rectangular member on top. In Xkipché, some of the cornice moldings consist of four members (illustrated in figure 4, bottom right). Other moldings only have one or two members.

4. Grammar-based Reconstruction

4.1. CGA Shape

CGA Shape is a grammar suitable for architectural design. In this subsection we will give a short introduction to CGA Shape necessary to encode the designs in the previous section. A more comprehensive description of CGA Shape is given in [MWH*06].

Shape: The grammar works with a configuration of shapes: a shape consists of a symbol (string), geometry (geometric attributes) and numeric attributes. Shapes are identified by their symbols which is either a terminal symbol $\in \Sigma$, or a non-terminal symbol $\in V$. The corresponding shapes are called terminal shapes and non-terminal shapes. The most important geometric attributes are the position P, three orthogonal vectors X, Y, and Z, describing a coordinate system, and a size vector S. These attributes define an oriented bounding box in space called *scope* (see figure 5).

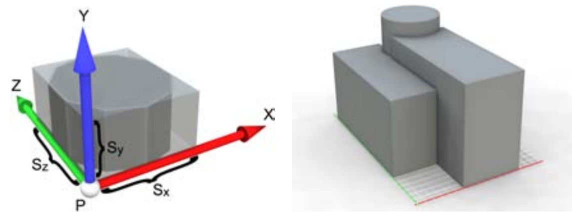


Figure 5: *Left: The scope of a shape. The point P, together with the three axis X, Y, and Z and a size S define a box in space that contains the shape. Right: A simple building mass model composed of three shape primitives.*

Production process: A configuration is a finite set of basic shapes. The production process can start with an arbitrary configuration of shapes A, called the axiom, and proceeds as follows: (1) Select an active shape with symbol B in the set (2) choose a production rule with B on the left hand side to compute a successor for B, a new set of shapes $BNEW$ (3) mark the shape B as inactive and add the shapes $BNEW$ to the configuration and continue with step (1). When the configuration contains no more non-terminals, the production process terminates.

Notation: The CGA Shape production rules are defined in the following form:

id: predecessor : condition $\rightsquigarrow$ successor

where *id* is a unique identifier for the rule, *predecessor* $\in V$ is a symbol identifying a shape that is to be replaced with *successor*, and *condition* is a guard (logical expression) that has to evaluate to true in order for the rule to be applied. For example, the rule

1: fac(h) : h > 9 $\rightsquigarrow$ floor(h/3) floor(h/3) floor(h/3)

replaces the shape fac with three shapes $floor$, if the parameter h is greater than 9. To specify the successor shapes we use different forms of rules explained in the remainder of this section.

Shape operations: Similar to L-systems we use general rules to modify shapes: $T(t_x, t_y, t_z)$ is a translation vector that is added to the scope position P, $Rx(angle)$, $Ry(angle)$, and $Rz(angle)$ rotate the respective axis of the coordinate system, and $S(s_x, s_y, s_z)$ sets the size of the scope. We use [and] to push and pop the current scope on a stack. Any non-terminal symbol $\in V$ in the rule will be created with the current scope. Similarly, the command $I(objId)$ adds an instance of a geometric primitive with identifier $objId$. Typical objects include a cube, a quad, and a cylinder, but any three-dimensional model can be used. The example below illustrates the design of the mass model depicted in figure 5 right:

1: A $\rightsquigarrow$ [T(0,0,6) S(8,10,18) I("cube")]
T(6,0,0) S(7,13,18) I("cube")
T(0,0,16) S(8,15,8) I("cylinder")

Basic split rule: The basic split rule splits the current scope along one axis. For example, consider the rule to split the façade of figure 6 left into four floors and one ledge:

1: fac $\rightsquigarrow$
Subdiv("Y",3.5,0.3,3,3,3){ floor | ledge | floor | floor | floor }

where the first parameter describes the split axis ("X", "Y", or "Z") and the remaining parameters describe the split sizes. Between the delimiter { and } a list of shapes is given, separated by |.

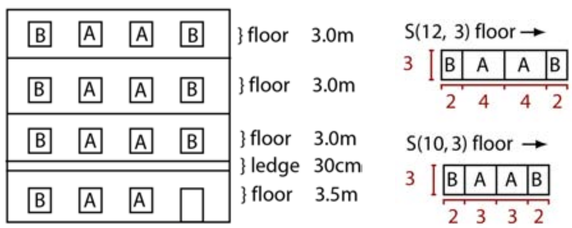


Figure 6: *Left: A basic façade design. Right: A simple split that could be used for the top three floors.*

Scaling of rules: From the previous example we can see the first challenge. The split is dimensioned to work well with a scope of size $y = 12.8$, but for other scopes the rule has to be scaled. From our experience not all architectural parts scale equally well, and it is important to have the possibility to distinguish between absolute values (values that do not scale) and relative values (values that do scale). Values are considered absolute by default and we will use the letter *r* to denote relative values, e.g.

1: floor ⤳ Subdiv("X",2,1r,1r,2){ B | A | A | B }

where relative values r_i are substituted as $r_i * (Scope.sx - \sum abs_i)/\sum r_i$ ($Scope.sx$ represents the size of the x-length of the current scope). Figure 6 right illustrates the application of the rule above on two different sized floors (with x-length 12 and 10).

Repeat split: To allow for larger scale changes in the split rules, we often want to tile a specified element. For example

1: floor ⤳ Repeat("X",2){ B }

tiles the floor into as many elements of type *B* along the x-axis of the scope as there is space. The number of repetitions is computed as $repetitions = \lceil Scope.sx/2 \rceil$ and we adjust the actual size of the element accordingly.

Component split: Up until this point all shapes (scopes) have been three-dimensional. The following command allows to split into shapes of lesser dimensions:

1: a ⤳ Comp(type, param){ A | B | ... | Z }

where *type* identifies the type of the component split with associated parameters *param* (if any). For example we write $Comp(\text{"}faces\text{"})\{A\}$ to create a shape with symbol *A* for each face of the original three-dimensional shape. Similarly we use $Comp(\text{"}edges\text{"})\{B\}$ and $Comp(\text{"}vertices\text{"})\{C\}$ to split into edges and vertices respectively. To access only selected components we use commands such as $Comp(\text{"}edge\text{"},3)\{A\}$ to create a shape *A* aligned with the third edge of the model or $Comp(\text{"}sidefaces\text{"})\{B\}$ to access all the side faces of e.g. a cube or polygonal cylinder.

4.2. Reconstruction Rules

In the following we explain the rules that we created to model Puuc buildings as they have been introduced in section 3. The rules are slightly simplified for space reasons, however, they are in principal sufficient to capture most of the variety of the Puuc architecture. The rules use *control parameters* that can be read from the GIS database. As geometric parameter we use the building footprint. The other control parameters are scalar values typeset in *italics*. The first rule takes the GIS footprint and extrudes it to a volumetric shape. Rule #2 creates façade shapes for each building face. As seen in figure 4 a façade can be broken down into several elements (rule #3). Note that we shorten height, width, depth or angle to *h*, *w*, *d* or *a* to save space in the rule description.

1: footprint ⤳ S(1r,*building_h*,1r) facades
2: facades ⤳ Comp("sidefaces"){ facade }
3: facade ⤳ Subdiv("Y",*base_molding*,1r, *medial_molding_h*, *frieze_h*,*cornice_molding_h*)
{ base_molding | lower_facade | medial_molding | frieze | cornice_molding }

This example shows a building with one door. Buildings with multiple doors can be generated by variations of rule #4. The buildings have only one façade with a door. All the other façades consist of a plain wall. The rule parameter *projection* in rule #6 is used to vary the thickness of the wall (by extruding the wall with distance *projection*, used for example in rule #16). The doorframe consists of three elements: the lintel on top and two jamb elements. Note that the lintel is arranged directly below the medial molding and that the jambs can be a little bit sloped (defined by angle given with the control parameter *jamb_a*).

4: lower_facade : Shape.visible("front") ⤳
Subdiv("X",1r,*door_w*+2**door_frame_w*,1r)
{ wall(0) | door | wall(0) }
5: lower_facade ⤳ wall(0)
6: wall(projection) ⤳ T(0,0,-projection)
S(1r,1r,*wall_d*+projection) I("wall.obj")
7: door ⤳
Subdiv("Y",*door_h*,1r){ Subdiv("X",1r,*door_w*,1r)
{ jamb(-*jamb_a*) | null | jamb(*jamb_a*) } | lintel }
8: lintel ⤳ S(1r,1r,*wall_d*) I("lintel.obj")
9: jamb(angle) ⤳ R(0,0,angle) S(1r,1r,*wall_d*) I("jamb.obj")

The moldings come in many varieties, but are all composed of the same elements. Figure 7 shows how the maximal four members of a molding are put together. The most important geometrical parameters are the heights of the elements and the projection parameters (to control how far elements are extruded from the plane that contains the wall).

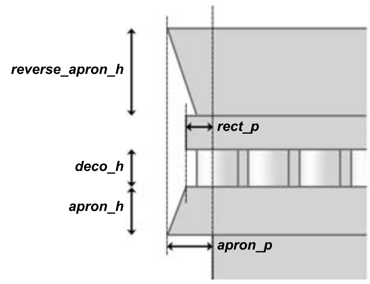


Figure 7: *The five parameters defining the appearance of a molding (plus the molding height). By setting the height of a member to zero the element can be switched off.*

All three different moldings (*base_molding*, *medial_molding* and *cornice_molding* as introduced in rule #3) are generated using the same procedure (initiated via rule #13). The only difference is the assignment of the different control parameters to the corresponding rule parameters:

10: cornice_molding ⤳ molding(*cornice_apron_h*, *cornice_deco_h*,*cornice_reverse_apron_h*, *cornice_apron_p*,*cornice_rect_p*)

This rule induces the creation of the cornice molding. The rules for the medial and base molding are defined accordingly, i.e. with the corresponding control parameters. In the following, we show the rule for a decorated molding. We also developed a similar rule to create moldings without decoration element (triggered via condition and then using other projection proportions).

13: molding(apron_h,deco_h,revap_h,apron_p,rect_p) $\rightsquigarrow$
Subdiv("Y",apron_h,1r,revap_h)
{ apron(rect_p,apron_p) | deco(rect_p*0.8-apron_p*0.2) |
rect(rect_p) | revap(rect_p*0.8-apron_p*0.2,apron_p) }

The four molding members are encoded as follows. All members have in common that they consist of a projected wall (described in rule #6): both aprons consist of a projected wall and a sloped element in front of the wall, the decoration member consists of colonettes in front of the projected wall, and the rectangular member consists of a projected wall only. Note that we call the projection parameters p, $p1$ and $p2$.

14: apron(p1,p2) $\rightsquigarrow$ wall(p1)
T(0,0,-p2) S(1r,1r,p2-p1) I("apron.obj")
15: deco(p) $\rightsquigarrow$ wall(p-*molding_colonette_diameter*/2)
T(0,0,-p) colonettes(*molding_colonette_diameter*)
16: rect(p) $\rightsquigarrow$ wall(p)
17: revap(p1,p2) $\rightsquigarrow$ wall(p1)
T(0,0,-p2) S(1r,1r,p2-p1) I("reverse_apron.obj")

The colonettes are created using cylinders. The rule just creates one empty element at the end so that the cylinders at the corner are not created twice. The modeling of building corners is a common challenge to most procedural building models.

18: colonettes(d) $\rightsquigarrow$ Subdiv("X",1r,d*1.2)
{ Repeat("X",d*1.2){ colonette(d) } | ε }
19: colonette(d) $\rightsquigarrow$ S(d,1r,d) I("cylinder.obj")

In the following we describe two of several different frieze types: one without and one with decoration (colonettes). Other frieze rules include the positioning of the masks (edge mask and front mask) which are separately reconstructed ornaments of high geometrical complexity.

20: frieze : *frieze_decoration* == "none" $\rightsquigarrow$ wall(*frieze_p*)
21: frieze : *frieze_decoration* == "colonettes" $\rightsquigarrow$
wall(*frieze_p*) colonettes(*frieze_colonette_diameter*)

5. Results

By using the rule set presented above, we are able to generate each of the building types listed in figure 4 in about 5 to 10 minutes (by simple modification of the control parameters). Additionally, we use rules for specifying materials and textures that are not shown in the text. Please note that we created all types according to archaeological data and we do not use random variations for the reconstruction. See figure 8 for selected buildings, with interesting height and molding combinations. These images are rendered in OpenGL and are screenshots from the interactive previewing system of the CityEngine. High quality renderings can be created with offline rendering.

Additionally, we extended the rules to generate other more complex buildings (few of them in Xkipché). Figure 9 shows a closeup of ornamented colonettes including mosaics and figure 10 pictures a whole building. These images have been created with Pixar's RenderMan. Ambient occlusion has been used to simulate the exterior lighting.

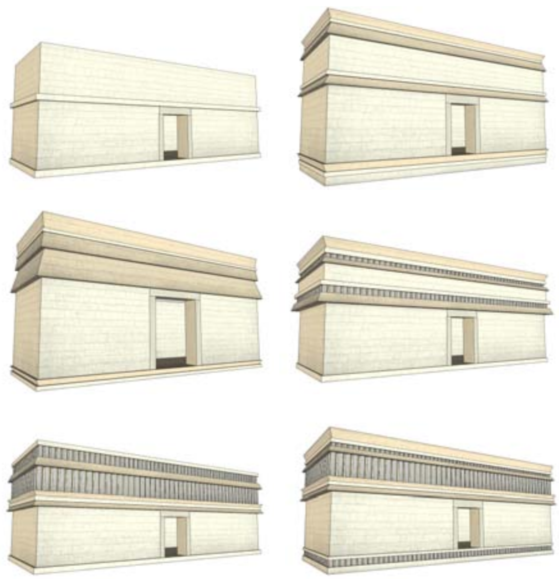


Figure 8: *This image shows various buildings that have been created in minutes by using the rule set described in the paper. Simple modifications of its control parameters lead to the different building appearances.*

Figure 9: *The rules for colonettes and frieze have been extended to be able to reconstruct also more complex building appearances.*

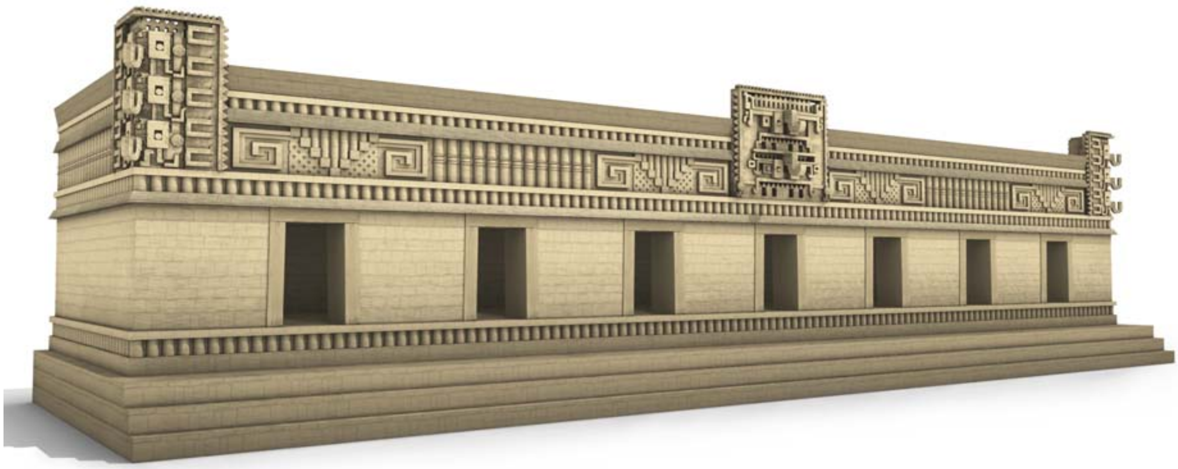


Figure 10: *Detailed reconstruction of one of the few highly ornamented buildings in Xkipché. The whole building has been generated procedurally by using the CityEngine, except the complex mask ornaments which have been created with traditional mesh modeling software. The image has been rendered in Pixar's RenderMan.*

6. Discussion

In this section, we want to identify contributions and open problems that are of interest for future research.

Procedural modeling and archaeology: Archaeology is an interesting application area for procedural modeling because information is only available in fragments. Therefore, the virtual reconstruction can not only be based on scanning, but needs to rely on human synthesis of data from multiple and heterogeneous sources. We believe that procedural modeling rules are an interesting and useful form of knowledge representation for such a synthesis. First, these rules can be used to create reconstructions that form the basis of archaeological discussion and presentation. Second, the rules themselves are formal and comprehensive. This notation has some advantages over a mere description in words, illustrated examples and annotated plans. The grammar framework ensures that essential parameters are not left unspecified and all information for a reconstruction is available.

Reconstruction detail: While we were able to create fairly detailed reconstructions of Xkipché buildings, modeling is generally an open ended problem and there are many opportunities for extensions: (1) It would be possible to integrate more GIS data, such as the exact position of the doors - now we only estimate door positions through the grammar. (2) Moldings could be made more accurate by using more parameters. (3) Additional (molding) decoration styles could be implemented - we discussed only the colonettes in detail. (4) Originally, the building surfaces were covered with plaster and painted with mineral and organic pigments [Car86], but since their exact appearance is still under archaeological debate, we did not include colorful textures such as paintings in the current model.

Efficiency and useability: The main part of the reconstruction work was reading and ordering the archaeological information and references (1 week). The model presented in this paper has then been created in three days: one day of architectural analysis, one day of modeling the elements (mainly the frieze-decoration was taking time) and one day for the actual implementation and encoding of the rules. Any building in Xkipché can now be reconstructed in detail in 2 minutes, if the user specifies the 20 - 30 parameters describing the building. We estimate that a professional CAD modeler will need 2-3h for the same task, but he would also need the models of the individual elements. Another big advantage of our approach is that after the initial model is created, archaeological researchers can create new models without any CAD-knowledge using a high-level user interface to specify parameters.

Future work: We are planning to investigate approaches to reconstruct the traditional houses (made of organic materials) and other aspects of urban environments, such as walkways and the vegetation in general. Therefore, we are working on a tighter GIS integration of the CityEngine by developing a practical GIS format which allows archaeologists to define the needed attributes directly in a common editor like ESRI's ArcGIS system. We expect that such an integration will enormously enhance the usefulness of our approach and make it applicable to all kinds of reconstruction scenarios.

7. Conclusion

We presented a method to procedurally create 3D reconstructions of stone houses in Xkipché. The reconstruction is based on archaeological research and makes use of shape grammars to encode the architectural design of buildings of the Puuc architecture. We demonstrated that this approach is a promising tool for archaeology, as it allows for the precise encoding of archaeological knowledge, simple and fast parameter-based modeling, and accurate 3D reconstructions of architectural content.

Acknowledgments

The authors thank Simon Haegler for helping with the renderings and the anonymous reviewers for their constructive comments on improving this paper. This research is supported in part by EC IST Network of Excellence EPOCH, EC IST Project CyberWalk, CHIRON Marie Curie EST Research Programme, NSF contract IIS-0612269 and NGA grant HM1582-05-1-2004. The excavation at Xkipché has been funded by the German Research Foundation (DFG): Archäologisches Projekt Xkipché (1991-1997) and Epiklassikum in Nord-Yucatan (2002-2006).

References

[And75] ANDREWS G. F.: *Maya Cities. Placemaking and Urbanization*. University of Oklahoma Press, 1975.

[BBJ*01] BIRCH P., BROWNE S., JENNINGS V., DAY A., ARNOLD D.: Rapid Procedural-Modelling of Architectural Structures. In *Virtual Reality, Archaeology and Cultural Heritage (VAST)* (2001), pp. 187–196.

[Car86] CARVER N. F.: *Silent Cities of Mexico and the Maya*. Documan Press Ltd, 1986.

[Dua02] DUARTE J.: *Malagueira Grammar – towards a tool for customizing Alvaro Siza's mass houses at Malagueira*. PhD thesis, MIT School of Architecture and Planning, 2002.

[GBHF05] GERTH B., BERNDT R., HAVEMANN S., FELLNER D. W.: 3D Modeling for Non-Expert Users with the Castle Construction Kit v0.5. In *Virtual Reality, Archaeology and Cultural Heritage (VAST)* (2005), Mudge M., Ryan N., Scopignooutit R., (Eds.).

[Hav05] HAVEMANN S.: *Generative Mesh Modeling*. PhD thesis, TU Braunschweig, 2005.

[LWH*05] LIU H., WANG Q., HUA W., ZHOU D., BAO H.: Building Chinese Ancient Architectures in Seconds. In *International Conference on Computational Science* (2005), pp. 248–255.

[MVUG05] MÜLLER P., VEREENOOGHE T., ULMER A., GOOL L. V.: Automatic Reconstruction of Roman Housing Architecture. In *International Workshop on Recording, Modeling and Visualization of Cultural Heritage* (2005), Balkema Publishers (Taylor & Francis group), pp. 287–297.

[MWH*06] MÜLLER P., WONKA P., HAEGLER S., ULMER A., VAN GOOL L.: Procedural Modeling of Buildings. *Proceedings of ACM SIGGRAPH 2006 / ACM Transactions on Graphics 25*, 3 (2006), 614–623.

[PL91] PRUSINKIEWICZ P., LINDENMAYER A.: *The Algorithmic Beauty of Plants*. Springer Verlag, 1991.

[PM01] PARISH Y. I. H., MÜLLER P.: Procedural modeling of cities. In *Proceedings of ACM SIGGRAPH 2001* (2001), Fiume E., (Ed.), ACM Press, pp. 301–308.

[Pol80] POLLOCK H. E. D.: *The Puuc. An Architectural Survey of the Hill Country of Yucatan and Northern Campeche, Mexico*. Peabody Museum of Archaeology and Ethnology Harvard University, 1980.

[Pre99] PREM H. J.: Geschichte eines Mayapalastes - Ausgrabungen in Xkipché, Yucatán. *Antike Welt 30*, 6 (1999), 545–554.

[Sti75] STINY G.: *Pictorial and Formal Aspects of Shape and Shape Grammars*. Birkhauser Verlag, Basel, 1975.

[Sti80] STINY G.: Introduction to shape and shape grammars. *Environment and Planning B 7* (1980), 343–361.

[WWSR03] WONKA P., WIMMER M., SILLION F., RIBARSKY W.: Instant architecture. *Proceedings of ACM SIGGRAPH 2003 / ACM Transactions on Graphics 22*, 3 (2003), 669–677.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Multi-Spectral Laser Scanning for Inspection of Building Surfaces
— state of the art and future concepts —

A. Wehr[1], M. Hemmleb[2] and Chr. Maierhofer[2]

[1]Institut for Navigation, University Stuttgart, Germany
[2]Federal Institute for Materials Research and Testing (BAM), Berlin, Germany

Abstract

The Federal Institute for Materials Research and Testing (BAM) developed a multi-spectral laser scanner to demonstrate the advantages of such systems for the inspection of building surfaces. It is shown that damage of building surfaces, caused by enhanced moisture content and/or vegetation, can be recorded automatically with a high signal-to-noise ratio by using four continuous wave semiconductor lasers at different wavelengths for the defined illumination of the object surface. It is worked out that the damaged areas can be identified by applying commercial multi-spectral image processing software. Finally a concept is presented to improve the overall performance of the system with regard to sampling speed and sensitivity. Here the cw lasers are modulated by sinusoidal signals of different frequencies and the intensity of the backscattered laser light is detected by synchronous demodulation. Furthermore, these structured signals can be used for slant range measurements. Knowing the slant for each pixel, three dimensional multi-spectral images are obtained.

Categories and Subject Descriptors (according to ACM CCS): I.4.1 [Digitization and Image Capture]: Scanning

1. Introduction

Precedent examination of building structures plays a crucial role for the planning of construction and restoration tasks. The multitude of possible damage of building surfaces makes manual damage mapping extensive and time-consuming. Therefore it is desirable to apply a method, based on image recording and processing for damage assessment, which also covers inaccessible parts of the building surface. It should also provide data for further analysis and classification of the existing damage. This approach offers good preconditions for automated damage detection. The surface structure and condition and especially its time-dependent changes have to be detected. Important damage of building surfaces are weathering, corrosion, salt blooming and biological changes like moss, lichen, moulds and moisture. The quantitative measurement of the moisture content has a particular importance, as most of the other damage is correlated with it.

Multi-spectral image recording techniques offer the possibility to detect damages on building surfaces. Image regions can be assigned to affected (e.g. damaged by moisture or vegetation) surface regions. The dimension of the damaged area can be measured quantitatively. Therefore, according to [BR82], [SSG90], [Fri03], [GAG95], [Ler01], [LRB00] and [Wig02] the results of a multi-spectral analysis especially in the near infrared (NIR) can serve as a basis for damage mapping, delivering quantitative results on the damaged area and about the nature of the damage.

Today multi-spectral imaging systems e.g. RGB cameras, infrared cameras and opto-electronical scanning devices all have in common that they are passive sensors which need external illumination. If sunlight is used, an illumination source with a large spectral bandwidth is available. As sunlight is changing with daytime season expensive calibration is required. In case of artificial illumination sources e.g. flash light or lamps one has to regard that all required spectral bands are existent. To circumvent the disadvantages of passive sensors BAM decided to develop an active multi-spectral sensor. As lasers offer well defined monochromatic optical radiation with the required intensity for remote sensing tasks and as they are small and robust, an imaging laser scanner using four semiconductor laser diodes working at different wavelengths was realized. The experiments carried out with this new device demonstrated that the recorded data could be processed with different commercial software tools enabling multi-spectral classification. It will be shown that information about the surface conditions and surface damages are gained from the difference in spectral reflectivity. As the spectral radiation powers of the lasers are known, even quantitative measurements of the damages are possible.

The sensitivity of BAM's multi-spectral laser scanning system can be improved, if the intensity of laser light is modulated with a known signal, e.g. a sinusoidal one, and detected by a correlation receiver. Applying this receiving technique the detected signal is not disturbed by background radiation e.g. extensive sunlight. Furthermore the modulation can be exploited for ranging. Besides actual temperature and humidity the slant range information is most important for quantitative measurements, because the intensity value for each pixel must be corrected by the free space loss which is a function of the instantanuous range.

2. Information Possibilities of Multi-Spectral Data

According to [Kra96] imaging spectrometers are applied to detect, measure and analyse the spectral content of the incident electromagnetic field besides the geometrical image content. The spectral information is required either to determine the chemical composition, the type and the physical condition of an object. After the acquisition imaging spectrometers deliver k images where k is the number of spectral bands available. In case of laser scanners number k is determined by the number of different lasers included in the scanner and determines the spectral resolution of the system. The spatial resolution is determined by the number of pixels within the field of view (FOV). The pixel size is given by the instantaneous field of view (IFOV). Figure 1 makes clear that after acquisition images with k layers are available in accordance to the k spectral bands. Figure 1 depicts that each pixel of one image layer contains the intensity of the received electromagnetic signal in the narrow spectral band of the corresponding sensor and laser respectively. Evaluating the intensities of all layers for a certain pixel, results in a characteristic spectrum for this pixel (s. Figure 1). This intensity spectrum represents the reflectance spectrum.

Each material and each sort of vegetation has a characteristic spectral reflectance (s. Figure 2) which can be used for computer aided classification. The classification is carried out in the feature space (s. Figure 3).

The studies carried out at BAM [HWSM06] make clear that commercial image processing software for remote sensing data can be applied on the gathered data sets. Best results were obtained by programs working with object based classification.

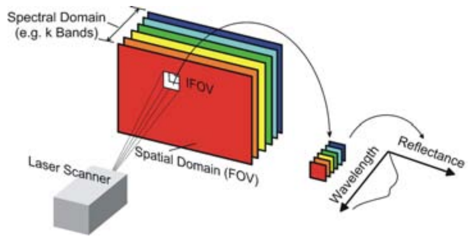


Figure 1: *Imaging spectrometer*

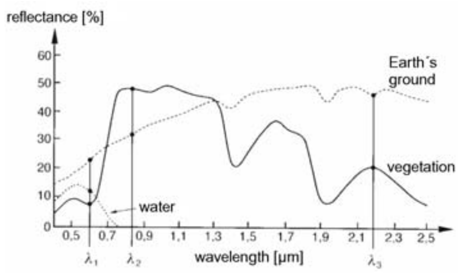


Figure 2: *Typical spectral reflectance of different materials (acc. [Alb01])*

In the following it will be worked out that vegetation indices can be computed from the laser scanner surveys by combining the different spectral layers also called image channels. The Normalized Difference Vegetation Index (NDVI) is calculated by

$$ndvi = \frac{x_{IR} - x_R}{x_{IR} + x_R} \qquad (1)$$

with x_{IR}: intensity measurement value at λ_{IR} = 808 nm
x_R: intensity measurement value at λ_R = 670 nm

It relates the intensity values measured in near infrared (x_{IR}) to intensity values obtained in red spectrum (x_R). For example infrared radiation is highly backscattered from the surface of healthy vegetation because chlorophyll exhibit a minimum of absorbtion in this spectral band. Whereas, the

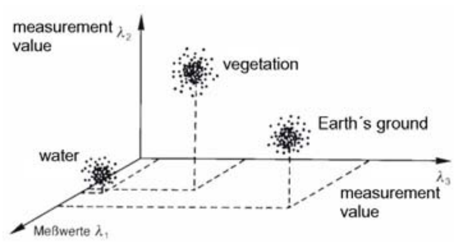


Figure 3: *Feature based space for multi-spectral classification (acc. [Alb01])*

red color laser beam shows low reflectance, because it is well absorbed. Exploiting these effects and determining NDVI vegetation on building surfaces can be detected easily and its distribution and by that the possible damage of a building can be classified.

According to NDVI the Normalized Difference Moisture Index (NDMI) was derived:

$$ndmi = \frac{x_{ref} - x_{IR}}{x_{ref} + x_{IR}} \quad (2)$$

with x_{ref} = intensity measurement value at reference wavelength, here at λ_{ref} = 980 nm
x_{IR} = intensity measurement value at λ_{IR} = 1930 nm

At this the measured intensity values spectrally located within the absorbtion line of water (here 1930 nm) are related to the intensity values obtained at reference wavelength (here 980 nm) which is insensitive against water contents. This means, NDMI represents a measure concerning the moisture content of the surface for the object under test.

The computation of NDVI and NDMI is integrated in the control and measurement software of the laser scanner described in the following chapter. After the measurement the user has available both indices as separate images, which permit an instant overview about the distribution of wet and vegetation at the building surface.

3. State of the Art of Multi-Spectral Laser Scanner Development

BAM built-up a breadboard multi-spectral laser scanner in the laboratory to analyse and to demonstrate the possibilities and limitation of such devices. The carried out experiments also offered design parameters for future multi-spectral laser scanners.

3.1. Multi-Spectral Laser Scanner

The laser scanner developed at BAM consists of four fibre-coupled semi-conductor laser diodes with different wavelengths, selected for the detection of moisture, natural cover and mineral changes. Considering the requirements for damage detection and the availability of laser diodes, the following wavelengths were chosen: 670 nm, 808 nm, 980 nm and 1930 nm. Suitable detectors based on Si and InGaAs photodiodes in combination with optical filters are used for the measurement of the reflected radiation.

For the realisation of a scanning measurement system, two different devices were selected and tested under laboratory conditions: a pan-tilt-unit as carrier for the multi-spectral-head (as shown in Figure 4 and Figure 5) and a mirror-scanning device, as known from 3D laser scanners. Geometric and radiometric calibration experiments were carried out.

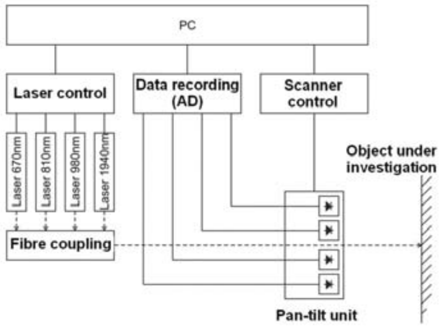


Figure 4: *Set-up of the multi-spectral laser scanner with pan-tilt unit*

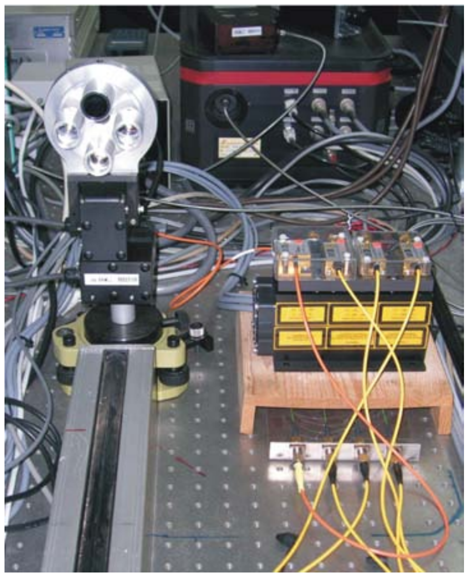


Figure 5: *Left: Pan-tilt unit with measurement head, right: laser diodes and fibre coupling unit*

Measurement control and data recording is run on a notebook. The results are saved in a data file containing the intensity of the detected reflections for the different wavelengths for each pixel as well as the 2D geometric position. For data analysis, the reflection intensity of each channel, the NDVI and the NDMI are plotted automatically as a function of geometric position. Additionally, the data can be saved in an output file, which can be directly used for enhanced multi-spectral classification with commercial software. The technical data are compiled in Table 1.

The first prototype of the multi-spectral laser scanner was set-up for laboratory investigation, and can principally be applied on-site, too. Further improvement is required for the robustness of the system, for the optimisation of the distance between measurement system and surface under investiga-

Table 1: *Key components of laser scanner*

laser	ALS multi-spectral laser system with semiconductor laser diodes with 670 nm, 808 nm, 980 nm, 1930 nm mounted on cooling radiator KB40
laser control	ALS (special design on basis of DioPower DP5)
optics	SuK fiber collimators (special design)
	Receiving optic for PTU (inhouse development)
	FhG IZM fiber couplers in different designs (special design)
detector	JI 577 Si photodiode with filter 670 nm, 810 nm, 980 nm J18 InGaAs photodiode with filter 1940 nm
data logger	NI DAQPad-6015 (USB)
scanner	DP PTU C46-17.5W with control PT-CB46C14
	Polytec mirror scanner OFV 040 with control DX 2102
controler	Sony Notebook 3GHz with software LabVIEW 7.1
visualisation and processing	LaserImage (inhouse development) Erdas Imagine 8.7

tion and for the measurement time per pixel, which can be drastically reduced by using the modulated laser light, e.g. pulsed or sinusoidal modulation.

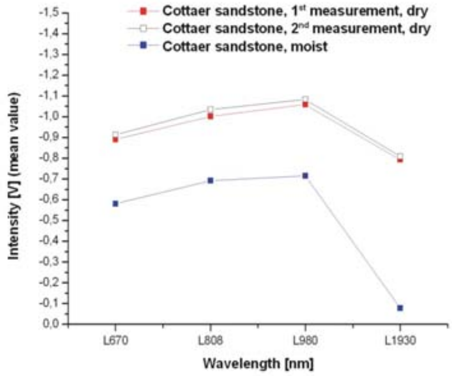


Figure 6: *Intensity of reflection for dry and moist sand stone at different wavelengths. The wavelength axis is not scaled*

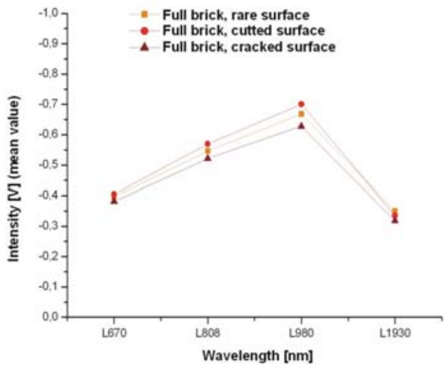


Figure 7: *Intensity of reflection at different surfaces (roughness). The wavelength axis is not scaled*

3.2. Measurement Results and Data Processing

Before BAM's multi-spectral laser scanner could be used for measurement tasks, test and calibration measurements had to be carried out. First, the radiometrical and geometrical reproducibility and stability of the whole system was tested and approved. This included the determination of the limits and the required parameters for later application. These measurement data were obtained by measurements of reflection properties of several different building materials (especially historic bricks and stones). During the tests material parameters such as moisture content and roughness of surfaces had been varied (see Figure 6 and Figure 7). For optimising the measurement parameters, the required power of the laser diodes and the distances of the scanner unit with respect to the measured surface were investigated systematically.

A small test specimen (wall consisting of several small cut bricks [cubes with a size of 5 x 5 x 5 cm^3]) was studied as shown in Figure 8. First, all bricks were measured in dry condition. Afterwards, two bricks were stored in water and the measurements were repeated for the determination of the moisture distribution at the surface. In addition to the four image channels for the four different wavelengths, also the NDMI was calculated from the reference wavelength (980 nm) and from the wavelength related to water absorption (1930 nm) as described above.

The distribution of the NDMI is displayed in Figure 9 as a false colour image. The area with enhanced moisture is clearly shown. At the bottom of the two moist bricks, water penetrated the joining dry bricks. Additionally, the differences in contrast due to varying surface properties recognised in Figure 8 outside the moist areas are not shown any more in Figure 9. The remaining small differences at the position of the edges of the bricks in Figure 9 and the increased NDMI at the bottom of the image are related to direct reflections of the laser radiation at the edges and to double reflections at the bearing surface of the test specimen, respectively.

The NDMI given in Figure 9 is only a measure for the relative moisture distribution at the surface of the specimen. For the determination of the absolute moisture content, a calibration of the measurement system with a reference measurement method is required.

Figure 8: *Wall consisting of small brick, two of the bricks are moist. The photo was taken with a digital camera*

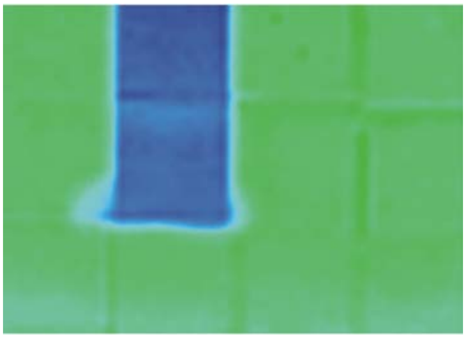


Figure 9: *Distribution of the NDMI calculated from the multi-spectral scan and being a measure for the relative moisture distribution*

4. Future Concepts for Multi-Spectral Laser Scanning

BAM's laser scanner transmits unmodulated signals. Therefore the backscattered laser signal from the object surface cannot be separated from background light. If the intensity of the laser radiation is modulated by e.g. sinusoidal signals the received backscattered laser light can be detected without disturbing background light. Those systems require correlation receivers which are also known as lock-in receivers. Using such technique improves the sensitivity and by that the range and possible measurement rate respectively. However, intensity modulation of laser light is a well established technique for 3D laser scanners using continuous wave semiconductor lasers [HW97], [FM04], [BV04]. [HW97], [FM04], [BV04] showed that high scanning rates with slant range accuracy down to sub-millimeter are possible. Therefore by integrating modulation techniques the multi-spectral laser scanner offers the possiblility to obtain 3D multispectral measurement data. 3D information is most important for the succeeding automatic measurement evaluation. The 3D data can be used to model the surveyed object surface. Knowing the surface topography the actual incident angel for each measurement point can be calculated. This improves the accuracy of the reflectance measurement. The exact knowledge of how the optical signal is damped by the atmoshpere is inalienable for measuring moisture parameters on building surfaces by infra red light. To regard precisely this effect the actual travelling length of the transmit and received laser radiation must be known. Then the free-space-loss and the influence of the atmosphere can be modeled.

In the following chapters the performance of 3D-cw-laser scanners will be outlined with regard to geometrical and intensity measurements. A demonstration experiment will depict that multispectral 3D-laser scanners can be realized with todays technology and will offer additional information, which can be processed by commercial remote sensing and classification software.

4.1. Intensity Modulation and Ranging

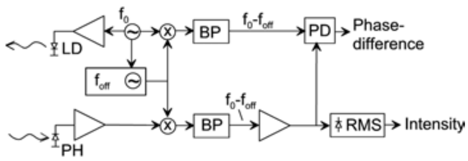


Figure 10: *Principle blockchart for synchronous demodulation*

Figure 10 depicts the blockchart for a typical transceiver with one signal using synchronous demodulation technique. The laser signal is intensity modulated with frequency f_0. The optical photodetector PH demodulates the frequency f_0. f_0 is down converted to the intermediate frequency f_0-f_{off}. Only this special frequency is able to pass the bandpass filter. Therefore, only this signal can be detected and the intensity can be measured. This process is also known as heterodyning. If the receiving branch is also down converted to f_0-f_{off} the phase difference between transmit and received signal can be measured. From the phase difference ϕ at intermediate frequency f_0-f_{off} the slant range r can be computed by

$$r = \frac{1}{4\pi}\frac{c}{f_0}\phi \tag{3}$$

if c is the speed of light and f_0 the frequency of the intensity modulation of the laser.

For a multispectral system with n lasers the system must be extended to n transceiver units. At the first glance such systems seem to be very expensive. However Figure 11 and Figure 12 demonstrate the gain in sensitivity and possible range performance. Both images were sampled with the same laser scanner in Museum for Natural Science Humboldt University Berlin. In Figure 11 the modulated laserlight was detected by synchronous demodulation. An avarage laser power of 5 mW was transmitted. The lower image shows the result with energy detection as applied by the BAM's multi-spectral laser scanner. In this case only the roof windows can be recognized, because they are illuminated by sunlight.

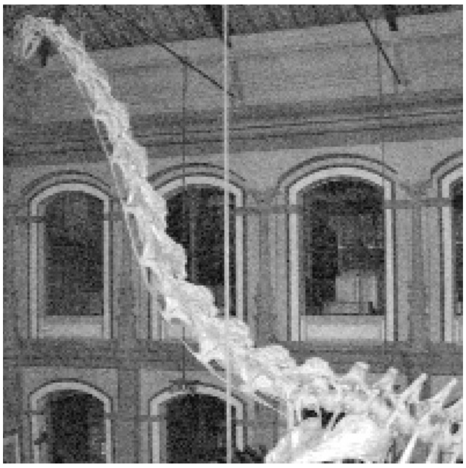


Figure 11: *Detection of the modulated laser signal*

Figure 12: *Energy detection of the unmodulated laser signal*

This experiment makes clear that the sensitivity can be improved by several orders of magnitude. Using modulated laser light also ranging can be carried out by measuring the phase difference between the transmitted and received signal. This measurement principle is realized in the 3D-Laser Scanner (3D-LS) of the INS [HW97]. The technical data are compiled in Table 2.

Table 2: *Technical parameters of 3D-LS*

laser power	0.5 mW
optical wavelength	670 nm / 830 nm
inst. field of view IFOV	0.1^{o}
total field of view FOV	30^{o}x 30^{o}
scanning pattern	2-dimensional line (standard) vertical line scan free programmable pattern
number of pixels per image	max. 32768 x 32768 pixels
range	<10 m, unambiguous range 15 m
ranging accuracy	0.1 mm (for diffuse reflecting targets with ρ=60% in 1m distance
side tone frequencies	10 MHz, 314 MHz
measurement rate	2 kHz (using on side tone) 600 Hz (using two side tones)

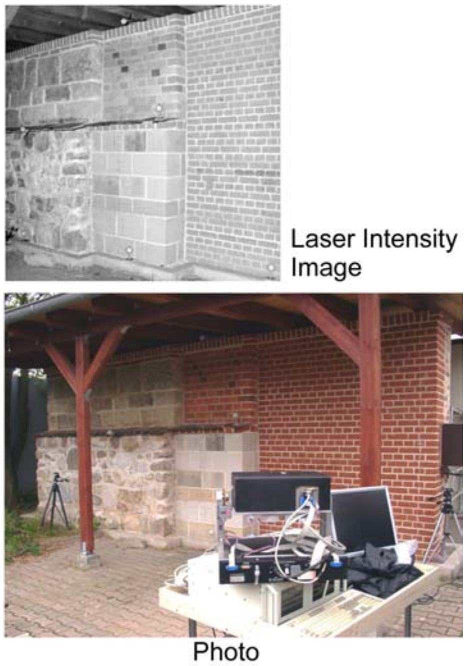


Figure 13: *Obelix: Measurement setup (photograph, lower image) and laser intensity image (upper image)*

Besides others the system was applied to survey the reference brickwork at BAM called "Obelix". It is a large masonry specimen built-up by using traditional historic construction methods and materials, e.g. several regionally typical materials like bricks (manufactured at the brick kiln in Glindow), sandstone, granite, limestone and bog iron stone. The intensity image demonstrates the photo quality (s. Fig-

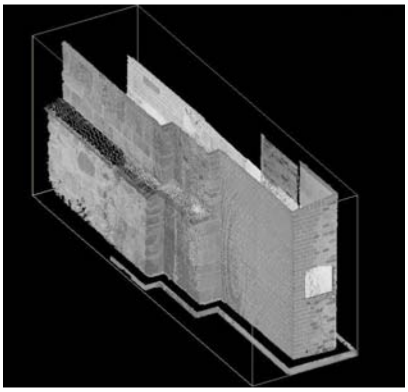


Figure 14: *3D-Animation view of Obelix 3D-LS data*

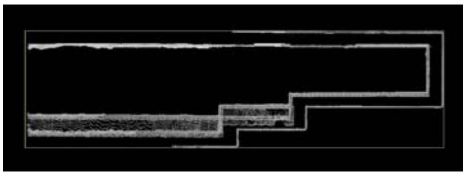


Figure 15: *Top-view of Obelix*

ure 13) of the system and the three dimensional performance can be learned from Figure 14 and Figure 15. This surveying demonstrates that the combination of 3D-information and intensity is most important, because the discrimination between the different materials cannot be carried out by the 3D-survey. Additional intensity information is required for a robust classification.

4.2. Possibilities of Multispectral 3D-LS

To study the possibilities of multi-spectral 3D laser scanners (3D-LSms) a laboratory scene with a banana plant was scanned with 3D-LS using semiconductor lasers at 670 nm and 830 nm. Figure 16 shows the surveying result. The plant exhibits high stress areas which can be well identified and classified (s. Figure 16). Figure 17 depicts the laser scanner survey result in ortho projection. The upper image is the composite of the spectral lines 670 nm and 830 nm. The decomposition of the two colours is shown in the lower images. The healthy parts of the plants appear very dark at 670 nm (red). This means the reflectance is very low. The stress areas can be recognized very well, because they exhibit a higher reflectance. In the NIR the total plant shows a higher reflectance. The contrast between stress and healthy areas is not as significant as at 670 nm. However the leaves are digitized comprehensively. This experiment demonstrates impressively the advantages of multispectral laser scanning. It makes clear that already with these two spectral laser lines a robust classification is possible and even a vegetation index can be calculated. Also three dimensional digitization is improved with regard to robustness, because if a certain spectral laser line is poorly scattered back from a target the others will compensate this effect.

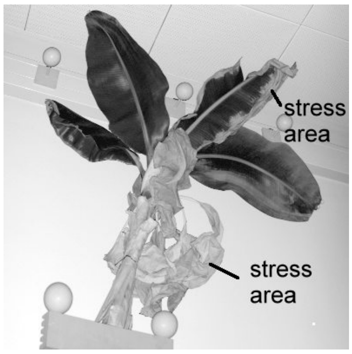


Figure 16: *Banana plant laser intensity image at 670 nm*

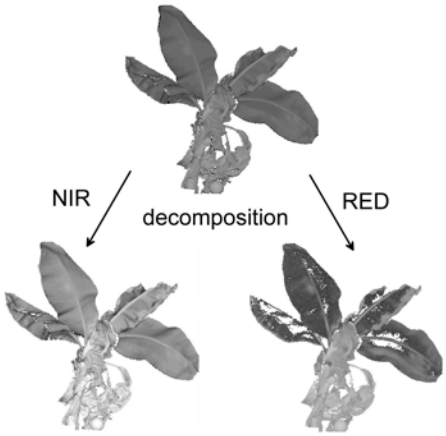


Figure 17: *Multispectral laser data (ortho projection)*

5. Conclusions

The evaluation experiments with BAM's multispectral laser scanner and the 3D-LS approve that imaging multi-spectral 3D laser scanning is a promising technology and certain classification tasks can be carried out with commercially available software. It was worked out very impressively that the sensitivity could be improved by orders of magnitude if a modulated laser signal is applied. A high sensitivity is required to reduce the transmitted laser power down to eyesafe levels and to cover ranges required for surveying of buildings.

In the application field of cultural heritage often questions are raised concerning the condition of historical masonry which is stressed by moisture and polluted by biological covering. Applying 3D-LSms, these loads on buildings can be detected by multispectral analysis, and the 3D-information is used for three dimensional modelling of the studied surface which allows a robust and quantifiable classification. For example regarding moisture measurements the total length of the laser link from measurement device to object surface and back and the impact direction with regard to surface normal must be known to obtain valid or even calibrated moisture measurements.

Looking at cultural heritage already a lot of surveying data are gathered in the interior of ancient housings, castles and caves. Here the optimum illumination is the major problem. As 3D-LSms is an active sensor, this problem will not exist anymore. Therefore, it is planned to develop a 3D-LSms for operational use on basis of the presented technology.

Acknowlegdements

The research project "Application of a multi-spectral laser scanner for the investigation of building surfaces" was funded by the Federal Office for Building and Regional Planning (BBR) (No. Z 6 - 10.07.03 / II 13 - 80010310). The authors would like to thank Dr. Friederike Weritz and Dieter Schaurich (BAM division VIII.2) for their support to this project, Anne Grote for programming control software, Gudrun Brinke for carrying out laboratory measurements, and Dipl.-Ing.(FH) Martin Thomas (INS) for carrying out laser scanner measurements and laser scanner data visualization.

References

[Alb01] ALBERTZ J.: *Einführung in die Fernerkundung*, 2 ed. Wissenschaftliche Buchgesellschaft Darmstadt, 2001.

[BR82] BÖTTCHER B., RICHTER H.: Ein Beitrag zum Nachweis von Feuchtigkeit in Mauerwerk mit Hilfe einer Infrarotkamera. *Materialprüfung 24*, 1 (1982), 5–9.

[BV04] BECKER R., VOLZ H.: IQSUN 880 a new modular concept for 3d-laser scanning. In *Laser-Scanners for Forest and Landscape Assessment* (Lemmer, Netherlands, 2004), Thies M., Koch B., Spiecker H., Weinacker H., (Eds.), vol. XXXVI, Part 8/W2, GITC, pp. 7–13. Proceedings of the ISPRS working group VIII/2 in Freiburg, Germany.

[FM04] FRÖHLICH C., METTENLEITER M.: Terrestrial laser scanning - new perspectives in 3d surveying. In *Laser-Scanners for Forest and Landscape Assessment* (Lemmer, Netherlands, 2004), Thies M., Koch B., Spiecker H., Weinacker H., (Eds.), vol. XXXVI, Part 8/W2, GITC, pp. 7–13. Proceedings of the ISPRS working group VIII/2 in Freiburg, Germany.

[Fri03] FRIEDMANN M.: Zerstörungsfreie flächenmässig erfassbare Ortung von Hohlstellen und Feuchtigkeitsbelastungen mittels Infrarot-Thermografie. *Schriftenreihe des Feuchte- und Altbausanierung e.V.*, 14 (2003), 61–70. 14. Hanseatische Sanierungstage.

[GAG95] GEYER E., ARNDT D., GÜNTHER B.: Bestimmung der Reflexion von trockenen und befeuchteten Baustoffen im Wellenlängenbereich 0,3 bis 2,5 μm für die Ir-optische Bestimmung der Feuchte. In *Feuchtetag 1995, BAM Berlin* (1995), pp. 176–184.

[GSG92] GODDING R., SACHER, G. SIEDLER G.: Einsatz von digitalen Aufnahmesystemen zur Gewinnung von Multispektralaufnahmen für die Erkundung von Bauwerksschäden. In *International Archives of Photogrammetry and Remote Sensing, Com. V* (New York, 1992), pp. 794–798.

[HW97] HUG C., WEHR A.: Applications of imaging laser sensors in close-range metrology and remote sensing. In *Optical 3-D Measurement Techniques IV* (Amsterdam, 1997), Wichmann, pp. 141–148. Conference in Zurich 29. Sep. - 2. Oct. 1997.

[HWS*06] HEMMLEB M., WERITZ F., SCHIEMENZ A., GROTE A., MAIERHOFER C.: Multi-spectral data acquisition and processing techniques for damage detection on building surfaces. In *ISPRS, Com. V Symposium, Dresden* (2006). 25.-27.9. 2006.

[HWSM06] HEMMLEB M., WERITZ F., SCHIEMENZ A., MAIERHOFER C.: Schadensdiagnose an Bauwerksoberflächen mit multispektralen Methoden. In *Bauwerksdiagnose 2006* (2006), vol. P08, pp. 277–286. DGZfp-Berichtsband 100-CD.

[HWWM04] HEMMLEB M., WERITZ F., WILSCH G., MAIERHOFER C.: Multispektraltechniken für die Schadensklassifizierung an Bauwerksoberflächen. In *Tagungsband zum Feuchtetag 2004 mit Workshop ON-SITEFORMASONRY* (2004).

[Kra96] KRAMER H.: *Observation of the Earth and Its Environment, Survey of Missions and Sensors*, 3 ed. Springer Verlag Berlin Heidelberg, 1996.

[Ler01] LERMA J.: Multiband versus multispectral supervised classification of architectural images. *Photogrammetric Record 17*, 97 (2001), 89–101.

[LRB00] LERMA J. L., RUIZ L. A., BUCHON F.: Application of spectral and textural classifications to recognize materials and damages on historic building facades. In *International Archives of Photogrammetry and Remote Sensing* (Amsterdam, 2000), vol. XXXIII, Part B5, pp. 480–483.

[SSG90] STRACKENBROCK B., SACHER G., GRUNICKE J.-M.: Image processing for mapping damages to buildings. In *Proceedings of the XIII. International CIPA symposium* (Cracow, 1990), p. 7.

[Wig02] WIGGENHAUSER H.: Active ir-applications in civil engineering. *Infrared Physics and Technology 43* (2002), 233–238.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Digital terrain modelling for archaeological interpretation within forested areas using full-waveform laserscanning

M. Doneus[1] and C. Briese[2]

[1]Department for Prehistory and Early History of the University of Vienna, Austria
[2]Christian Doppler Laboratory "Spatial Data from Laser Scanning and Remote Sensing",
Institute of Photogrammetry and Remote Sensing of the Vienna University of Technology, Austria

Abstract

The identification of sites within forested areas is one of the remaining unresolved issues for archaeological prospection. Airborne laser scanning can be a solution to this problem: due to the capability of penetrating forest to a certain degree (depending on the vegetation density) the determination of the terrain surface is even possible in wooded areas. To be able to identify archaeological structures, archaeologists have to interpret the resulting topographical data of a filtered ALS scan. This does not pose major problems with large structures. Smaller features, however, are much more difficult to identify, because their appearance in an ALS point cloud is very similar to natural and recent features, as for example dense brushwood, or piles of twigs or wood. Therefore, to eliminate potential sources of error, a high quality separation of terrain and off-terrain points is essential for archaeological interpretation while maintaining a high point density of the ALS data. Using conventional ALS systems, the possibilities to classify terrain and off-terrain points are limited and the results - especially in forested areas with dense understorey - are far from ideal for archaeological purposes. This paper will demonstrate how the new generation of full-waveform ALS systems can be used to get a much better classification of solid ground and vegetation cover and consequently DTMs, which can be interpreted archaeologically with much more confidence.

1. Introduction

Airborne Laser Scanning (ALS), also referred to as LIDAR (Light Detection and Ranging) [Ack99] [Kra04] [WL99] is an active remote sensing technique. It is used to produce dense and high precision measurements of the topography of the Earth's surface. The scanning device is typically mounted at the bottom or below an aeroplane or helicopter. For the determination of object points, the laserscanner emits short infrared pulses into different directions across the flight path towards the earth's surface and a photodiode records the backscattered echo and determines the distance to the reflecting object by the determination of the travel time.

Due to the capability of penetrating forest to a certain degree (depending on the vegetation density) the determination of the terrain surface is even possible in wooded areas [KP98] [PKK99]. However, in order to eliminate remaining off-terrain points from the derived last echo point cloud for the determination of the DTM advanced filter methods are necessary [SV03].

For archaeological interpretation purposes, a high quality separation of terrain and off-terrain points is essential. Using conventional ALS systems, the possibilities to classify terrain and off-terrain points are limited and the results - especially in forested areas with dense understorey - are far from ideal for archaeological purposes. This paper will demonstrate how the new generation of full-waveform ALS systems can be used to get a much better classification of solid ground and vegetation cover and consequently DTMs, which can be interpreted archaeologically with much more confidence.

2. Requirements of ALS for archaeological prospection

Though ALS has in the meanwhile found application in many areas, at present it finds only a very few archaeological applications related to archaeological prospection mainly in England and Germany [Bar03] [Bew03] [Cha06] [DACC05] [Hol01] [HHB02] [Mot01] [SR04] [Sit04].

With human occupation, the surface of the ground usually gets shaped (banks, ditches, mounds, etc.). When a site is abandoned, the structures will start collapsing. Without human interaction (especially agriculture), the process of decay can be stabilised by vegetation and the archaeological structures will form part of the ground's surface.

Because of its potential to determine a dense point cloud (up to several points per m^2) in high quality, archaeologists can investigate such sites, which are still surviving in relief even in forested areas. To be able to identify archaeological structures, archaeologists have to interpret the resulting topographical data of a filtered ALS scan. This does not pose major problems with large structures, like ruined castles, earthworks, tracks, or ridge and furrow. Smaller features, however, like round barrows, slag heaps, or kilns (to mention just a few) are much more difficult to identify. Their appearance in an ALS point cloud is very similar to natural and recent features, as for example dense brushwood, or piles of twigs or wood, which are - other than archaeological structures - actually off-terrain points.

Therefore, to eliminate potential sources of error, a high quality separation of terrain and off-terrain points is essential for archaeological interpretation while maintaining a high point density of the ALS data. Blunders and vegetation have to be filtered away, while the topography of buildings and any other man made structures (like ditches, banks, round barrows, walls etc.) should not be altered by the filtering techniques.

3. Problems of "Conventional" Data for Archaeological Interpretation

During the ALS point determination, the laser beam travels towards the earth's surface and illuminates different targets and all backscattering objects within the footprint area contribute to the received echo. The backscattering characteristics of the target depend on its size, distance, its reflectivity, and the directionality of scattering.

Flat surfaces in respect to the beam direction without vegetation reflect one short echo (approx. the same length as the emitted pulse), which the scanner receives as a delayed and attenuated signal. However, at spatially distributed targets, as it is typically the case in forested areas, only some of the laser energy is scattered back by the treetops, while the other energy penetrates through and is reflected by branches, bushes and the ground later on. In that way, a single laser beam can scatter back a complex echo waveform. Here, the returning signal is a superposition of echoes from the different scatterers. The echoes can be received as distinct signals if separated by distances larger than the range discrimination of the ALS system [WUD*06]. Within currently available commercial systems, this resolution is typically around 1.5 m [Kra04]. While this discrimination is high enough to clearly distinguish trees from the ground surface, it implicates that with those systems, it is difficult to distinguish narrow vegetation (as ferns, bushes etc.) from the terrain.

The elevation accuracy of ALS is composed of a systematic shift value (which mainly comes from the inaccuracies of GPS and INU measurements) and an unsystematic inaccuracy, which is mainly caused by the different range distribution within the footprint (e.g. vegetation cover of the ground [PKK99]). While the systematic shift value does not pose any problem for archaeological interpretation if the relative errors between two overlapping ALS strips are not too big, the unsystematic inaccuracies from the vegetation cover will result in topographic features representing for example single bushes, dense vegetation, or tree trunks. However, these off-terrain objects are sometimes difficult to discern from smaller archaeological features, as mentioned above.

Usually, for obtaining a DTM from a conventional scan, some of these topographical features can be sorted out by various filtering techniques. For archaeological purposes, however, filtering cannot be applied too rigid. Otherwise, smaller archaeological features are removed or smoothed, too. Therefore, while high vegetation and houses can be removed by filtering more or less reliably, low vegetation, like bushes and tree-trunks will still survive in the resulting DTM and will cause problems during interpretation.

4. Full-Waveform Sensors

The problems posed above mainly occur, because typical ALS sensors can only record up to a certain degree of distinct (typically 2 to 4) echoes from multiple targets touched by a single laser pulse using analogue detectors in real time during the acquisition process. As a result, these systems provide "only" an irregular 3D point cloud containing coordinates of the detected echoes.

However, the latest generation of commercially available ALS systems allow to discretisise the full-waveform of the received echo for each emitted laser beam. This discretisation allows us to gain further physical observations of the reflecting surface elements, which can be useful for a subsequent object classification. By modelling the full waveform as a series of Gaussian distribution functions, individual scatterers can be distinguished [HMB00] [WUD*06]. The results are estimates of the location and scattering properties of the individual targets: for each returning echo of a single laser pulse, the estimated coordinates of the scatterer, the echo width, and the amplitude is determined.

The echo width gives us information on the range distribution of all the small individual scatterers contributing to one echo, whereas the amplitude gives information about the radiometric scattering properties of the illuminated targets that contribute to one echo. If the echo width is small, a rather flat surface element was illuminated, whereas when it is large, scatters at different ranges contribute to the one determined echo (as it is the case with a tilted surface or a terrain surface with narrow vegetation (flowers, fern, small bushes)). In these cases, the estimated distance to the object

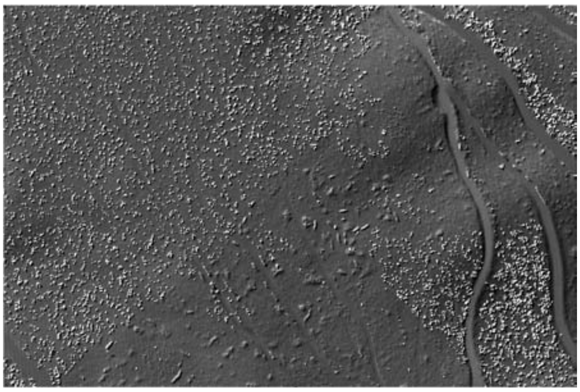


Figure 1: *DSM resulting from the unfiltered last echo point-cloud. There are many points, which represent, tree-trunks, very dense vegetation or narrow vegetation, which do not represent the actual ground.*

is a mixture of all distances to the individual targets. In the presence of low vegetation, the height of the estimated point will be too high and the point will not exactly represent the terrain surface.

The amplitude can give us additional information on the quality (intensity) of the reflection. It is, however, much more difficult to use the amplitude for classification purposes, because different effects as for example footprint area, the directionality of scattering, as well as size, topography, vertical distribution (number of backscatters from a single laser pulse), and reflectivity of the target contribute to a single amplitude. Therefore, it may not be as straightforward to use for classification purposes as reflectivity values recorded with passive imaging sensors [KK04].

Using amplitude and echo width, it is possible to investigate the return signal and extract additional ground characteristics. Consequently, much more information is available when classifying the point cloud into solid ground and vegetation cover.

5. Case Study

In April 2006, we launched the project "LiDAR-Supported Archaeological Prospection in Woodland", which is funded by the Austrian Science Fund (FWF P18674-G02). The goal of the project is to explore the potential of ALS for Archaeological Prospection in a densely forested area; specifically, to evaluate an approx. 190 km^2 forest area within the Leitha mountain range in an archaeological case study.

5.1. Test Data

During the pilot phase of the project, an ALS scan was conducted in the dormant period beginning of April 2006 using the latest generation of full-wave recording scanning systems. The purpose of this first flight was to get the optimal sensor configuration for scanning the whole area with its specific forest structure and canopy characteristics later on. The test area was carefully selected to represent different canopy density over already known archaeological features.

For the scans, we were using the RIEGL ALS-system LMS-Q560 operated by the company *Milan Flug GmbH*. The LMS-Q560 digitally samples and stores the entire echo waveform of the reflected laser pulses [HUG04]. Regarding the physical properties of ALS systems [Bal99], the LMS-Q560 has following specifications [WUM*04]: laser wavelength (1,5 μm), pulse duration (4 ns), pulse energy (8 μJ), pulserate (<100 kHz), beam width (0.5 mrad), scan angle (+/-22,5 deg), flying height (<1.500 m) and size (0.5 m @ 1 km) of the laser footprint on ground. Its multi-target range discrimination is 0,6 m.

The scan was performed on April 8, 2006. Flight altitude was about 600 m above ground, which resulted in a laser footprint size of 30 cm on ground. A total area of 9 km^2 was covered with a scan angle of +/-22,5 degrees by 26 parallel flight tracks, which had a width of approximately 500 m and an overlap of 50%. The real scan rate was 66 kHz that resulted in an overall mean point density of eight measurements per m^2. The GPS operated with a frequency of 1Hz, whereas the inertial measurement unit (IMU) recorded the attitude with 250Hz. While scanning, the total area was additionally covered vertically using the integrated DigiCAM H/22 system with a resolution of 22 megapixels.

5.2. Full-Waveform Processing and Filtering

In the first step a Gaussian decomposition of the full-waveform data was performed [WUD*06]. Based on this procedure the 3d co-ordinates of all detected echoes are determined and stored together with the additional information (amplitude, echo width) determined from the fitted Gaussians. From this dataset, we selected the last echoes together with the additional information for the further processing.

To eliminate remaining off-terrain points from the derived last echo point cloud (Figure 1), we used the theory of robust interpolation (RI) [KP98] within a hierarchic framework [BPD02], which is implemented into the software package SCOP++ [KO05]. The most important feature within this coarse to fine approach is the method of RI. RI integrates the elimination of off-terrain points and the interpolation of the terrain surface within the same process. The two most important entities of the RI are a functional model and a weight model. The functional model has to allow the approximation of the terrain surface with the consideration of individual weights for each irregularly distributed point. The weight model has to assign an individual weight to each point.

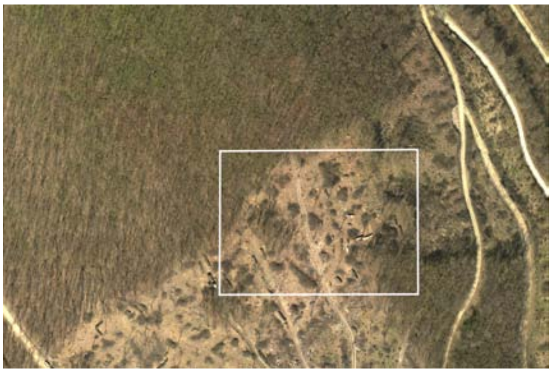


Figure 2: *Vertical aerial photograph of study area 1 taken during the scanning procedure. The white rectangle marks the outline of Figure 6.*

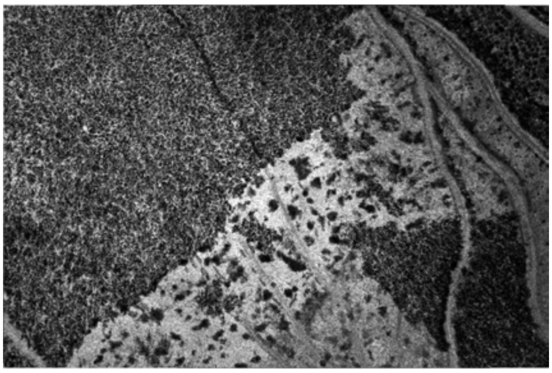


Figure 3: *Intensity image of study area 1. Bright pixels represent a high percentage of the original laser energy returned as last-pulse echo.*

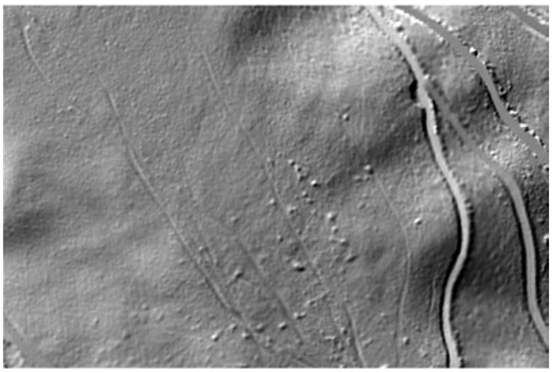


Figure 4: *DTM of study area 1 after Gaussian decomposition and filtering using robust interpolation with an eccentric and unsymmetrical weight function within a hierarchic framework.*

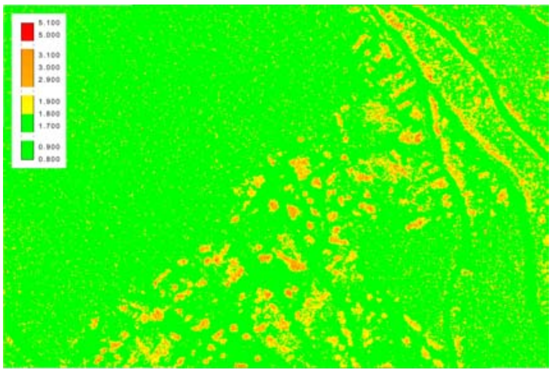


Figure 5: *Mapped echo widths of study area 1. While in the wood, high values (yellow to red) occur only occasionally, bushes and clearance piles show high values throughout.*

Figure 6: *Zoomed view of 3D point cloud and orthophotograph after application of the threshold value. Most clearance and even wood piles are filtered out and are not represented in the point data any more.*

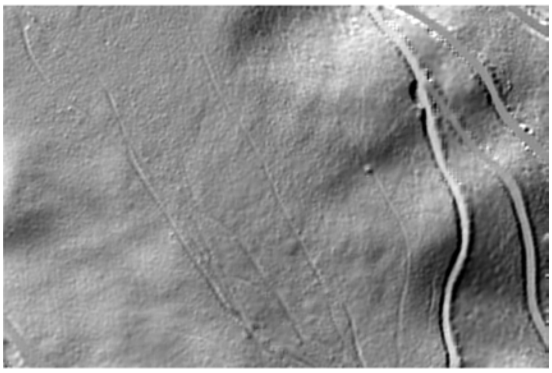


Figure 7: *Resulting shaded DTM after removal of points with a high echo width and hierarchic robust interpolation.*

For the elimination of off-terrain points from ALS data, an eccentric and unsymmetrical weight function is typically in use [BPD02]. This weight function assigns iteratively a low weight to points that are significantly above the terrain (e.g. on the vegetation) whereas terrain points are assigned a high weight. The RI is applied within each data pyramid level of the hierarchic setup. For the stepwise refinement, a tolerance band defined in respect to the resulting coarse surface model is used. This step-by-step process leads to a final surface model that allows to separate terrain from off-terrain points using a threshold value (typically plus and minus three times sigma of the height accuracy of the input points). The whole process of filtering is also demonstrated at the EuroSDR Distance Learning Course "Filtering and Classification of Laser Scanner Data", available under http://www.ipf.tuwien.ac.at/eurosdr/index.htm.

In an archaeological dataset, vegetation should be filtered out, but walls, banks, ditches, and smaller archaeological structures should be kept within the resulting DTM. One therefore has to be very careful with the tuning of the parameters of the iterative filtering steps within the hierarchic filter strategy. Due to the fact that no big buildings were present within our test area, we started the procedure with two data pyramid levels and did not use the additionally available pre-elimination of large building regions. For the first level, a point density of 1 point/m^2 was chosen. After the thin out of the original point cloud (approx. 8 points/m^2), we performed the RI using an asymmetric weight function in order to eliminate the off-terrain points within this data pyramid level. This was followed by the refinement step with the help of a tolerance band in respect to the 1 m-terrain model that discarded all last echo points 0.5 m above the terrain surface. Finally, we applied the RI to the remaining data of the finest level.

This procedure was applied on one hand to all last echo points without considering the additional information provided by the full-waveform processing, whereas on the other hand the same process was applied to a subset of the points after a simple pre-elimination step of all last echo points with a large echo width. To demonstrate the potential of this way of point classification and the improvement achieved with the consideration of the additional point attributes gathered by the full-waveform analysis, two small areas from the test-scan will be presented in the following.

5.3. Study Area 1

The first area does not contain archaeological structures. The vertical photograph (Figure 2) shows that the upper area is covered with deciduous wood, while in the lower area, the wood is being cleared. Apart from a few groups of trees, the lower area is covered with clearance piles of twigs and logs from the felled trees. On the right side, there is a bundle of roads, where dense bushes accompany the individual roads.

The situation is well reflected in the intensity image (Figure 3), which is derived from the amplitude values, which we got from Gaussian decomposition of the full-waveform data. The bright pixels show that a high percentage of the original laser energy returned as last-pulse echo. The darker the pixels get, the less energy returned to the sensor from the ground. Deciduous forest and small groups of trees consist of a mixture of high and low amplitude values, while the reflections coming from the cleared soil have high amplitudes. Clearance piles show equal areas of low amplitude values.

After Gaussian decomposition and filtering, the resulting shaded DTM shows that the procedure worked quite well in the wooded area (Figure 4). In the cleared area, the DTM shows still many small features with a diameter of 5 - 6 m and a height of 0,2 - 1 m. Without further information, it would be hard to interpret and decide about their potential archaeological meaning. Additionally, the dense bushes next to the roads are still present.

If we map the individual echo widths (Figure 5), we clearly can identify the piles of twigs and logs, which had low amplitude values as having also large echo widths. Analysing the mapped echo widths and their histogram, we defined a threshold value (in this case 1,7 m) and consequently removed all points from the original data set with higher values (Figure 6).

The resulting DTM shows a much better surface representation with almost all of the vegetation and clearance piles removed (Figure 7). This result is now much easier to interpret with a minimized risk of identifying low, dense vegetation as potential archaeological feature.

5.4. Study Area 2

In the other study area, we wanted to test, whether the procedure of eliminating points with higher pulse-width would affect archaeological features. Therefore, we analysed a small area within the Iron Age hillfort of Purbach. The area under investigation shows parts of the ramparts and a graveyard consisting of at about 50 round barrows (Figure 8).

The massive ramparts are clearly visible despite the vegetation in the aerial photograph (Figure 8). The round barrows, however, vanish under the dense vegetation cover. The intensity image (Figure 9) let us distinguish three different kinds of vegetation covering the barrows: bushes (1), trees without brushwood (2), and trees with brushwood (3).

After elimination of all echoes with a high echo width (threshold value: 1.7 m) (Figure 10), the comparison between the two DTMs (with and without pre-eliminated points) clearly shows, that the archaeological features were not affected by the procedure (Figure10 and Figure 11). The difference map reveals that the measurements of only a few barrows had been affected by low vegetation, which was removed consequently (Figure 12).

Figure 8: *Vertical aerial photograph of study area 2. While the massive ramparts are clearly visible, the round barrows (1) vanish under the dense vegetation cover.*

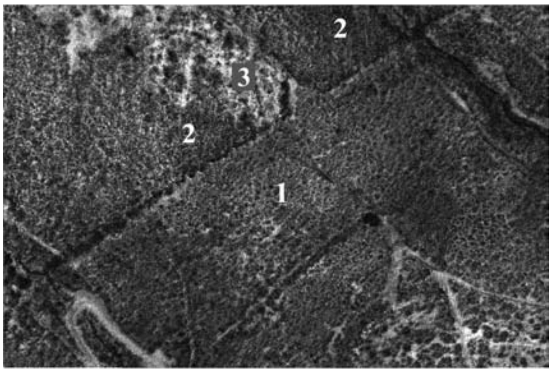


Figure 9: *Intensity image of study area 2; the amplitude values let us distinguish three different kinds of vegetation: bushes (1), trees with brushwood (2), and trees without brushwood (3).*

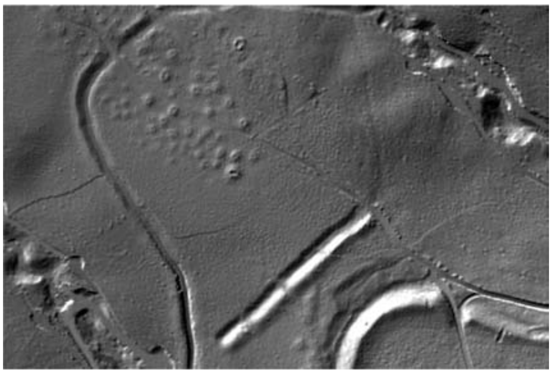


Figure 10: *Resulting shaded DTM after hierarchic robust interpolation.*

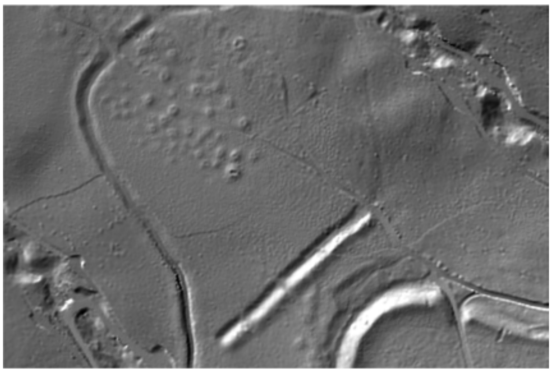


Figure 11: *Resulting shaded DTM after removal of points with a high echo width and hierarchic robust interpolation.*

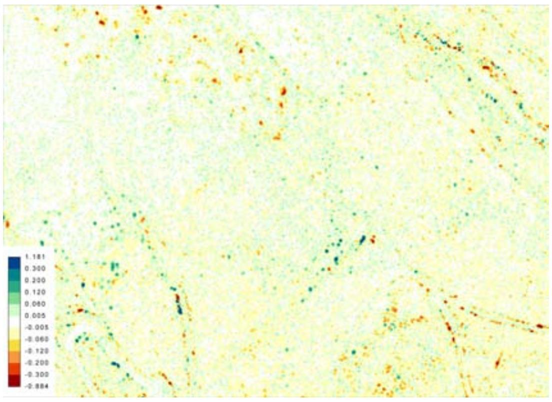


Figure 12: *Z-coded difference between DTM1 and DTM2 (with and without eliminated points).*

6. Conclusion

The paper demonstrated the potential of full-waveform analysis. With a full-waveform recording, a lot of more information about the received laser echo, like the detailed distribution of targets in the beam path, their reflectance and their vertical extent is stored.

In that way, it is possible to investigate the return signal and therefore further point characteristics can be determined. Based on this information, we can now classify the ALS points with much more confidence into solid ground and vegetation cover. This has an important impact (not only) on archaeological interpretation: in archaeological terms, we now have means to identify features in the point-cloud, which are remains of bushes or other low vegetation that was not completely penetrated by the laser pulse. After eliminating these points, we get a more reliable DTM, where most local topographic features are in fact local topographic elevations of the ground surface.

Within the paper, the usage of a simple threshold operation in order to pre-exclude points situated within low vegetation structures is demonstrated. However, this simple threshold procedure can be problematic in steep terrain, where the illuminated tilted terrain surface within the footprint can increase the echo width. Therefore, we will extend our procedure for steep terrain in the future. Furthermore, we will consider a calibration of the sensor system in order to determine a system independent interpretation of the received amplitude values [WUD*06]. This calibration procedure should allow to convert the received echo amplitude into the radar backscatter cross section which provides a system independent information about the radiometric surface reflection properties.

Acknowledgements

This research has been supported by the Austrian Science Fund (FWF) under project no. P18674-G02. The authors thank Thomas Melzer from the Christian Doppler Laboratory for Spatial Data from Laser Scanning and Remote Sensing, Vienna University of Technology, for the processing of the Gaussian decomposition of the full-waveform ALS data.

References

[Ack99] ACKERMANN F.: Airborne laser scanning: present status and future expectations. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 2-3 (1999), 64–67.

[Bal99] BALTSAVIAS E.: Airborne laser scanning: existing systems and firms and other resources. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 2-3 (1999), 164–198.

[Bar03] BARNES I.: Aerial remote-sensing techniques used in the management of archaeological monuments on the british army's salisbury plain training area, wiltshire, uk. *Archaeological Prospection 10* (2003), 83–91.

[Bew03] BEWLEY R.: Aerial survey for archaeology. *The Photogrammetric Record 18*, 104 (2003), 273–292.

[BPD02] BRIESE C., PFEIFER N., DORNINGER P.: Applications of the robust interpolation for dtm determination. In *Photogrammetric Computer Vision* (Graz, Austria, 2002), Kalliany R., Leberl F., Fraundorfer F., (Eds.), International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV, 3A, pp. 55–61.

[Cha06] CHALLIS K.: Airborne laser altimetry in alluviated landscapes. *Archaeological Prospection 13*, 2 (2006), 103–127.

[DACC05] DEVEREUX B., AMABLE G., CROW P., CLIFF A.: The potential of airborne lidar for detection of archaeological features under woodland canopies. *Antiquity 79* (2005), 648–660.

[HHB02] HOLDEN N., HORNE P., BEWLEY R.: High-resolution digital airborne mapping and archaeology. In *Aerial Archaeology Developing Future Practice*, Bewley R., Raczkowski W., (Eds.), Nato Science Series, Series I: Life and Behavioural Sciences, Vol. 337. 2002, pp. 173–180.

[HMB00] HOFTON M. A., MINSTER J. B., BLAIR J. B.: Decomposition of laser altimeter waveforms. *IEEE Transactions on Geoscience and Remote Sensing 38* (2000), 1989–1996.

[Hol01] HOLDEN N.: Digital airborne remote sensing: the principles of lidar and casi. *AARGNews 22* (2001), 23.

[HUG04] HUG C., ULLRICH A., GRIMM A.: Litemapper-5600, a waveform-digitizing lidar terrain and vegetation mapping system. In *Proceedings of Natscan, Laser-Scanners for Forest and Landscape Assessment - Instruments, Processing Methods and Applications* (2004), Thies M., Koch B., Spiecker H., Weinacker H., (Eds.), International Archives of Photogrammetry and Remote Sensing, Volume XXXVI, Part 8/W2, pp. 24–29.

[KK04] KATZENBEISSER R., KURZ S.: Airborne laser-scanning, ein vergleich mit terrestrischer vermessung und photogrammetrie. *Photogrammetrie, Fernerkundung, Geoinformation 3* (2004), 179–187.

[KO05] KRAUS K., OTEPKA J.: Dtm modelling and visualization - the scop approach. In *Photogrammetric Week* (Heidelberg, 2005), Wichmann Verlag, pp. 241–252.

[KP98] KRAUS K., PFEIFER N.: Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing 53*, 4 (1998), 193–203.

[Kra04] KRAUS K.: *Photogrammetrie. Band1. Geometrische Informationen aus Photographien und Laserscanneraufnahmen. 7., vollständig bearbeitete und er-*

weiterte Auflage. Walter de Gruyter, Berlin, New York, 2004.

[Mot01] MOTKIN D.: The assessment of lidar for archaeological use. *AARGNews 22* (2001), 24–25.

[PKK99] PFEIFER N., KRAUS K., KÖSTLI A.: Restitution of airborne laser scanner data in wooded areas. *GIS Geo-Information-Systems 12*, 2 (1999), 18–21.

[Sit04] SITTLER B.: Revealing historical landscapes by using airborne laser scanning. a 3-d modell of ridge and furrow in forests near rastatt, germany. In *Proceedings of Natscan, Laser-Scanners for Forest and Landscape Assessment - Instruments, Processing Methods and Applications* (2004), Thies M., Koch B., Spiecker H., Weinacker H., (Eds.), International Archives of Photogrammetry and Remote Sensing, Volume XXXVI, Part 8/W2, pp. 258–261.

[SR04] SHELL C., ROUGHLEY C.: Exploring the loughcrew landscape: a new airborne approach. *Archaeology Ireland 18/2*, 68 (2004), 22–25.

[SV03] SITHOLE G., VOSSELMAN G.: Comparison of filtering algorithms. In *Proceedings of the ISPRS working group III/3 workshop "3-D reconstruction from airborne laserscanner and InSAR data"* (Dresden, Germany, 2003), Maas H.-G., Vosselman G., Streilein A., (Eds.), International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV, 3/W13, pp. 71–78.

[WL99] WEHR A., LOHR U.: Airborne laser scanning - an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 2-3 (1999), 68–82.

[WUD*06] WAGNER W., ULLRICH A., DUCIC V., MELZER T., STUDNICKA N.: Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing 60*, 2 (2006), 100–112.

[WUM*04] WAGNER W., ULLRICH A., MELZER T., BRIESE C., KRAUS K.: From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges. In *Geo-Imagery Bridging Continents. XXth ISPRS Congress, 12-23 July 2004 Istanbul, Turkey* (Istanbul, 2004), Altan O., (Ed.), International Archives of Photogrammetry and Remote Sensing, Volume XXXV, Part B/3, pp. 201–206.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

COMBINING LASER SCANNING AND PHOTOGRAMMETRY - A HYBRID APPROACH FOR HERITAGE DOCUMENTATION

Norbert Haala, Yahya Alshawabkeh

Institute for Photogrammetry, Universitaet Stuttgart, Germany

Abstract

High quality 3D models of cultural heritage sites can be generated efficiently by laser scanning, which allows the accurate and dense measurement of surface geometry. In addition to the geometric data collection, texture mapping based on additionally collected digital imagery is particular important for this type of application. This requires a combined processing of range and image data sets. For this purpose, they have to be registered or aligned by a suitable transformation to a common reference coordinate system. The involved transformation parameters can be determined based on corresponding elements to be extracted from the different data sets. In the paper an efficient edge detection algorithm is presented, which allows for the automatic segmentation of such primitives even in complex scenes. In order to achieve a high quality 3D photo-realistic mode,l this alignment process has to be followed by an automatic texture mapping, which is discussed in the second part of the paper. The presented algorithms are demonstrated in the framework of a project aiming at the generation of a 3D virtual model of the Al-Khasneh, a well-known monument in Petra, and a Roman Theatre in ancient Jerash city, Jordan.

Categories and Subject Descriptors: I.4.8 [Scene Analysis]:Range data, I.4.7 [Feature Measurement]: Texture

1. Introduction

Terrestrial laser scanning is frequently used to provide high quality 3D models of cultural heritage sites and historical buildings. Based on the run-time of reflected light pulses, these sensor systems allow for the fast, reliable and area covering measurement of millions of 3D points. While this enables an effective and dense measurement of surface geometries, the provision of image data is frequently additionally required for a number of applications. This is especially true in the context of heritage documentation, since the complete documentation of heritage sites usually implies a high quality texture mapping based on supplementary image data. For this reason, some commercial 3D systems directly integrate a digital camera in order to simultaneously collect corresponding RGB values for each LIDAR point. This induces camera viewpoints, which are identical to the laser scanning stations. However, these camera viewpoints might not be optimal for the collection of high quality imagery as they are required for texture mapping. Additionally, laser scanning for the documentation of complex object structures and sites frequently has to be realised from multiple viewpoints. This can result in a relatively time consuming process. For outdoor applications these large time differences will cause varying conditions of illumination and changing shadows. Thus, the captured images will be subject to considerable radiometric differences, which will disturb the visual appearance of the resulting textured model. For these reasons, the acquisition of object geometry and texture by two independent sensors and processes will be advantageous. This allows for an image collection at optimal positions and time for texturing, which is especially important for the high requirements to be met during the realistic documentation and subsequent visualisation of heritage sites.

While an independent image collection considerably improves the quality of the available texture images, data processing and evaluation will complicate. If a camera is directly integrated with the laser scanning system, the captured images can be directly linked to the 3D point cloud, if a proper calibration of the complete system is available. In contrast, for independent camera and laser scanner stations, the combined evaluation requires a co-registration or alignment of the collected range and image data sets as a first processing step. This co-registration is usually realised based on corresponding primitives within the different data sets. For this purpose, primitives like 3D points or lines have to be extracted from the range data and matched against their 2D projections in the image. While a number of algorithms are available to determine the camera pose from such 2D-3D correspondences, the automatic provision of such point or line features is difficult. Within section 2 an algorithm, which can generate suitable primitives based on efficient edge detection also in complex scenes is presented.

If the camera pose is available, the texture images can be directly mapped to the corresponding 3D surface patches for visualisation. However, texture should of course only be assigned to those parts of the 3D object model, which are actually visible. This requires the detection of occlusions in the respective texture images. While this can be solved relatively easy for simple object geometries, the effort for visibility analyses increases considerably for geometrically complex object representations as they are usually provided from laser scanning in the context of cultural heritage documentation. In these scenarios the 3D object model is frequently represented by meshed surfaces. Since these meshes are derived from dense 3D point clouds, a huge amount of surface patches is generated. Thus, efficient approaches have to be made available to solve the visibility problem. As it will be discussed in section 3 this can be realised based on approaches originally developed in computer graphics.

2. Feature Extraction and Registration

Registration, or alignment, is the first step to be solved for the combined processing of range and image data collected from different viewpoints. Frequently, different laser scans are combined using measurements at artificial targets like spheres or signals, which can be detected and identified easily in the range data. Since such targets are also clearly visible in images, they can additionally be used for the registration of the image data. However, the provision and measurement of such targets requires additional effort. This is especially true, if a large number of potential texture images have to be collected. Additionally, these targets may occlude important parts of the object within the texture images [LHS00]. For these reasons, it is advantageous to use natural features to provide the required corresponding primitives for registration.

Traditionally, in photogrammetry corresponding points are used for determination of the image pose. As it is demonstrated exemplarily for a cultural heritage application in Figure 1 (left), suitable points can be identified well in an image. In contrast, the exact measurement of corresponding points within data from laser scanning is almost impossible. This is deomnstrated for the captured 3D point cloud (Figure 1 middle) as well as the range image (Figure 1 right), which was derived from laser measurements by re-interpolation to a regular grid. Due to the difficulties in exactly identifying such point structures, the measurement accuracy frequently is not sufficient for registration [LS05].

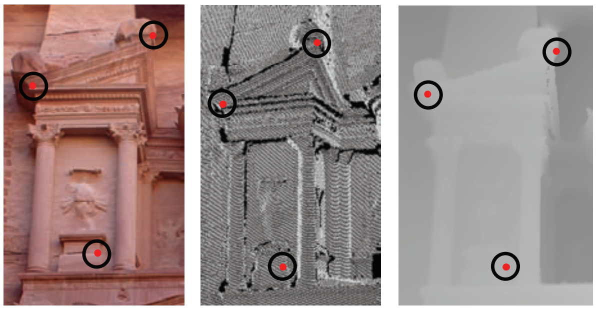


Figure 1: Corresponding points in image, point cloud and range image.

In contrast, linear features can usually be extracted and measured more accurately and reliably [CH99]. Such edge structures are frequently available within scenes of man made environments and allow for the reliable pose estimation by spatial resection. Thus, an automatic extraction of such edge structures is beneficial compared to a measurement of distinct points [LH02]. For this reason an efficient segmentation algorithm was developed to automatically extract such features of interest from the collected range and image measurements.

2.1. Curvature based range image segmentation

Usually, range data segmentation does not use unordered 3D point clouds (Figure 1 middle), but is based on 2.5D representations by raster grids (Figure 1 right). These range images maintain the original topology of the original laser scans. Thus, neighborhood relations are available implicitly and tools from image processing can be adopted. For this reason an easier implementation of segmentation algorithms is feasible compared to unordered 3D point clouds.

Similar to image segmentation, existing approaches for range data processing can be categorized in region-based and edge based techniques. Region based approaches group range pixels into connected regions by some homogeneity measure. A survey is e.g. given in [HJJ*96] and [MLM02].

While region based approaches allow for a reliable extraction of smooth or planar surface patches from range data, the direct segmentation of edges is difficult. This results from the limited spatial resolution of range data. Additional problems result from the large amount of noise at such height discontinuities, frequently occurring due to multipath effects. Thus, only a few edge based segmentation algorithms have been developed [SD01]. However, since edges are very well suited for co-registration of range and image data, this was our motivation for the development of an algorithm for the extraction of such structures from range images. There, the range image edges are extracted based on the analysis of the mean curvature values. The surface is approximated locally by an analytic representation, which is used to calculate the different properties of the respective patches.

By these means, the mean curvature at edges is used to detect local maxima or zero crossings in the range image. Within further processing steps, a multi-scale edge detection and a subsequent skeletonization is used to increase the reliability and accuracy of edge detection and localization.

2.2. Computation of Mean Curvature Values

Range data segmentation requires an appropriate surface description. This description should be rich, so that matches of similar elements can be detected, stable so that local changes do not radically alter the descriptions, and it should have a local support so that the visible objects can be easily identified. These characteristics are provided by the mathematical properties of the mean curvature, which is closely related to the first variation of a surface area. Unlike the Gaussian curvature, the mean curvature depends on the embedding, for instance, a cylinder and a plane are locally isometric but the mean curvature of a plane is zero while that for a cylinder is non-zero. Mean curvature is invariant to arbitrary rotations and translation of a surface, which is important for surface shape characterization. Since mean curvature is the average of the principal curvatures, it is slightly less sensitive to noise during numerical computations. Due to these characteristics, mean curvature values can provide stable and useful measures for detecting surface features in range and intensity images.

Several techniques are known for the efficient estimation of the mean curvature. As an example analytical methods fit a surface to a local neighbourhood of the point of interest. This surface approximation is then used to compute the partial derivatives needed to calculate the curvature values. Our approach is based on the work of [BJ88], who proposed an analytical technique for estimating the mean and Gaussian curvature. The advantage of this approach is its flexibility to estimate the curvature values at multiple scales, and the efficient computation of the values by optimized convolution operations. The approach and can be summarized as follows: For a given odd $N \times N$ window, each data point is associated with a position (u, v) from the set $U \times U$ where

$$U = \{-(N-1)/2, ..., -1, 0, 1, ..., (N-1)/2\}$$

The local biquadratic surface fitting capability is provided using the following discrete orthogonal polynomials:

$\O_0(u)=1$, $\O_1(u)=u$, $\O_2(u)=(u^2 - M(M+1)/3)$; $M=(N-1)/2$

To estimate the first and second partial derivatives, an orthogonal set of $d_i(u)$ functions using the normalized versions of the orthogonal polynomials $\O_i(u)$ is used:

$$\vec{d}_i(u) = \frac{\O_i(u)}{P_i(M)} : P_0(M) = N, P_1(M) = \frac{2}{3}M^3 + M^2 + \frac{1}{3}M.$$

$$P_2(M) = \frac{8}{45}M^5 + \frac{4}{9}M^4 + \frac{2}{9}M^2 - \frac{1}{9}M^2 - \frac{1}{15}M.$$

Since the discrete orthogonal quadratic polynomials over the 2D window are separable in u and v, partial derivative estimates can be computed using separable convolution operators. These derivatives estimates can then be plugged into the equation for mean curvature. The equally weighted least squares derivative estimation window operators are then given by:

$$[D_u] = \vec{d}_0 \vec{d}_1^T, [D_v] = \vec{d}_1 \vec{d}_0^T,$$

$$[D_{uu}] = \vec{d}_0 \vec{d}_2^T, [D_{vv}] = \vec{d}_2 \vec{d}_0^T, [D_{uv}] = \vec{d}_1 \vec{d}_1^T$$

$\tilde{g}(i,j)$ represents the noisy, quantized discretely sampled version of a piecewise-smooth graph surface. Afterwards, the partial derivative estimate images are computed via appropriate 2D image convolutions.

$$\tilde{g}_u(i,j) = D_u \otimes \tilde{g}(i,j), \tilde{g}_v(i,j) = D_v \otimes \tilde{g}(i,j),$$

$$\tilde{g}_{uu}(i,j) = D_{uu} \otimes \tilde{g}(i,j), \tilde{g}_{uv}(i,j) = D_{vv} \otimes \tilde{g}(i,j),$$

$$\tilde{g}_{uv}(i,j) = D_{uv} \otimes \tilde{g}(i,j)$$

The mean curvature is then computed using the partial derivatives estimates as the following:

$$H(i,j) = \frac{(1+\tilde{g}_v^2(i,j))\,\tilde{g}_{uu}(i,j) + (1+\tilde{g}_u^2(i,j))\tilde{g}_{vv}(i,j) - 2\tilde{g}_u(i,j)\,\tilde{g}_v(i,j)\tilde{g}_{uv}(i,j)}{2(\sqrt{1+\tilde{g}_u^2(i,j)+\tilde{g}_v^2(i,j)})^3}$$

2.3. Mean Curvature Analysis

The behaviour of the mean curvature for specific object properties can be demonstrated well by the filter results for synthetic range images. Thus the mean curvature was computed for range images of a block and a wye, which are depicted in Figure 2. The curvature values were then extracted at the horizontal profile represented by the line overlaid to the respective range image. From the analysis of these curvature values as they are depicted in the bottom of Figure 2 one can conclude that:

a) For jump edge boundaries (J) where surface depths are discontinuous, the mean curvature exhibits a zero crossing. Two distinct peaks of opposite algebraic sign are clearly visible in the profile of computed curvature values.
b) For crease edges (C) at discontinuities in the surface normal direction, the curvature response is a smooth peak. Concave (Cv) and convex (Cx) edges can be discriminated by the different algebraic sign of the curvature values. The exact position of a convex crease edge is defined by the maximum curvature value, while the concave crease edge is given at a minimum.

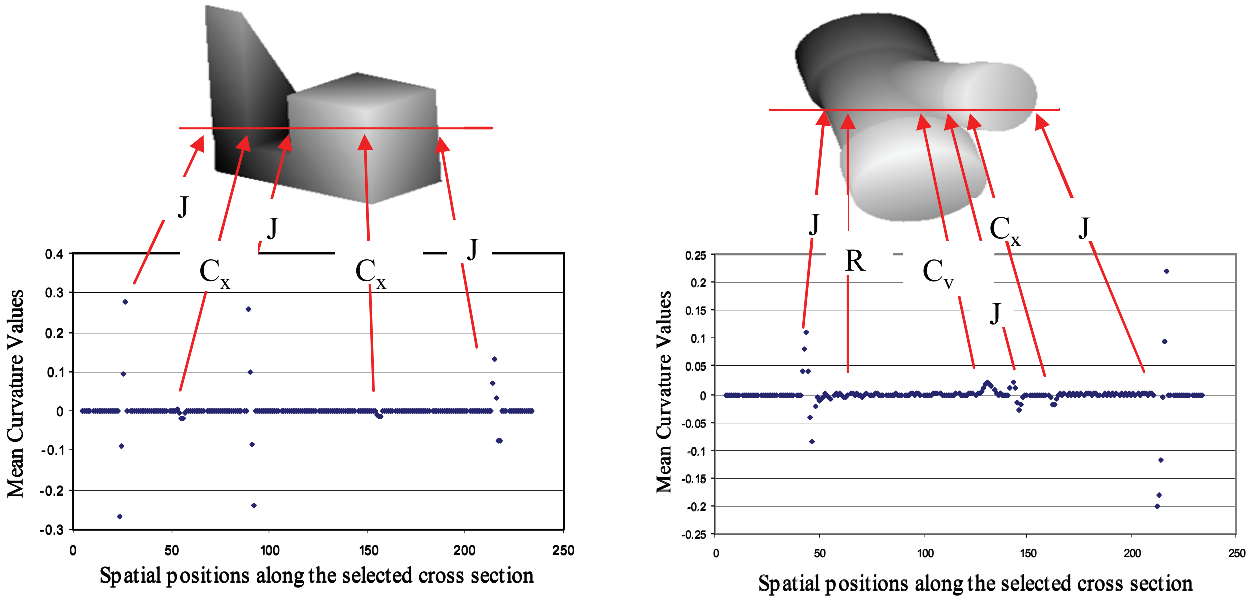


Figure 2: *Spatial distribution of the mean curvature values for block and wye range image*

c) At ridges (R) the mean curvature also indicates a change in the orientation of the surface normal, however, the response is smaller compared to crease edges.

d) Compared to crease edges, the values of the mean curvature are larger at jump edges. Their value mainly depends on the magnitude of the depth discontinuity.

e) For jump edges, the exact position is defined at a zero crossing between two peaks of the mean curvatures, whereas for both crease and ridge edges the true edge is given by the maximum and minimum value of the peaks.

After computation of the mean curvature values $H(x,y)$ a pixel represents an edge location $\{(x,y):E(x,y)=1\}$ if the value of the gradient exceeds some threshold. Thus:

$$E(x,y)=\begin{cases}1\ if\ \| H \| > T\ for\ some\ threshold\ T\\ 0\ otherwise\end{cases}$$

In order to locate the position of crease, ridge and step edges, zero crossings as well as smooth peak values are searched within the computed mean curvature values. Since the value of the mean curvature is smaller for crease edges than for jump edges, the edge types of an object can be classified easily by applying different threshold values. Low threshold values are used to detect the small peaks of crease edges while larger values can be used for step edge detection. This ability of our algorithm to reliably characterize these edge types is demonstrated exemplarily in Figure 3.

Figure 3: Curve block segmentation using different thresholds to detect step edges (red) and crease edges (blue).

Figure 4 displays the segmentation results for two range data sets for the 3D model of Al-Kahsneh monument in Petra, Jordan. The data were collected by a Mensi GS100 laser scanner [AH05]. As already mentioned, the segmentation is based on range images, which, in contrast to unordered 3D point clouds, maintain the original topology of a laser scan and thus allow an easier implementation of the segmentation algorithms. The top row of Figure 4 depicts the outer façade of Al-Kahsneh, while the bottom row shows data collected for one of the interior rooms. The second column, referred as Figure 4b shows the binary edge maps as they are generated using the curvature based segmentation. As it is visible, most of the main features are detected. Since a large filter mask size was used, the edges are rather blurred. For this reason, an edge thinning is applied. Figure 4c depicts the resulting skeletons overlaid to the original range image.

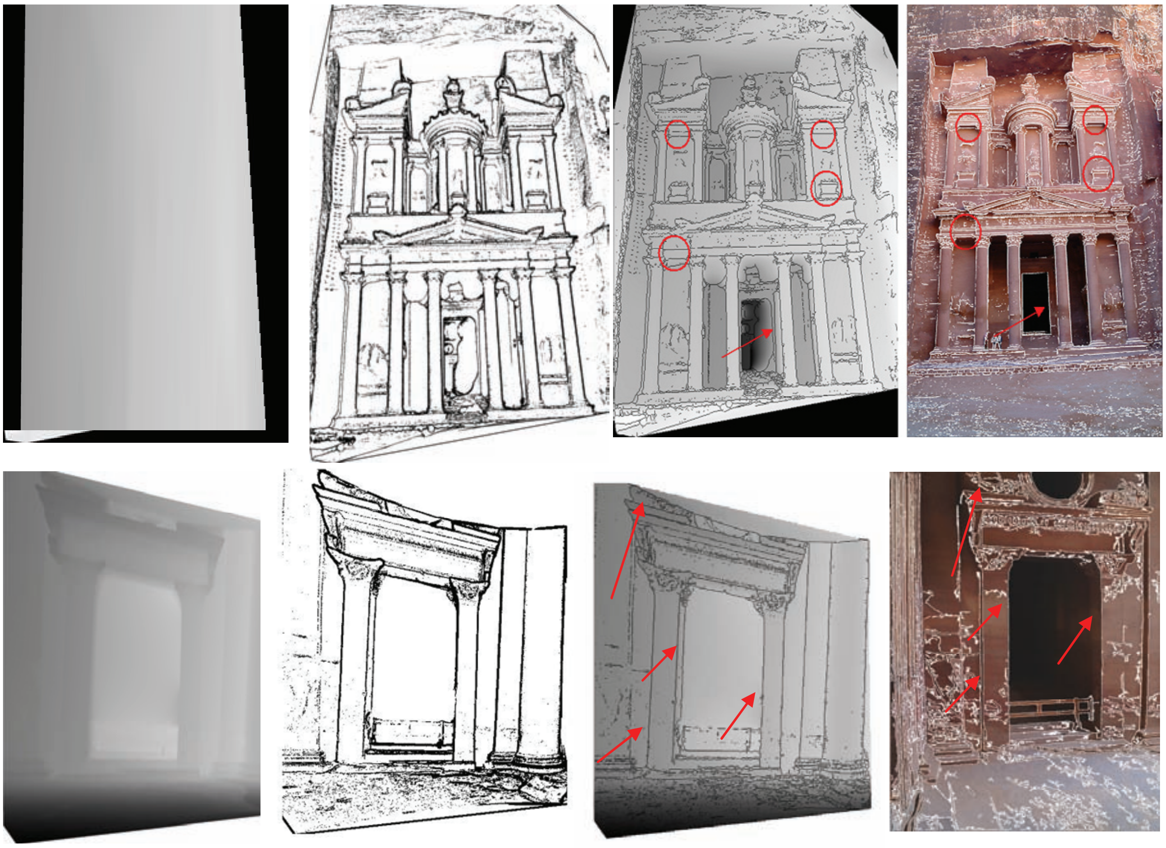


Figure 4: *a) Range images for 3D model of Al-Khasneh. b) Binary range edge images. c) Segmentation results after thinning overlaid to corresponding range image. d) Segmentation results of colored image. Red arrows and circles show selected edges used for registration of the 2D-3D data sets.*

As it is depicted in Figure 4d our segmentation is not limited to range images, but can also be applied for feature extraction from intensity images. In addition to the segmentation results, Figure 4c and Figure 4d also show some manually selected corresponding lines, which were used to register both colour and range images based on an algorithm developed by [KF03]. Thus, the linear features as they are extracted by our segmentation process allow for a precise and reliable co-registration of range and image data sets.

3. Multi-image texturing of complex 3D scenes

After this registration, texture mapping can be realized by warping the images onto the collected object surfaces. For this purpose, the transformation parameters as determined in the preceding step are used. Additionally, it has to be guaranteed that the image is only mapped onto parts of the 3D model that are actually visible from the respective camera viewpoints. Thus, a occlusion detection is additionally required before texture mapping. While such a visibility checking within the texture images can be realised without problems for relatively simple object representations, it can become very tedious and time consuming within the context of cultural heritage documentation. As an example, for documentation of the Al-Kasneh monument shown in Figure 4, laser scans were collected for three different viewpoints. After registration and meshing, this resulted in a 3D model at an average resolution of 5 cm with more than 2 million triangles. During texture mapping in principle every triangle node has to be mapped from object space to the corresponding point on the texture image and the visibility of each mesh in the respective images has to be checked.

Efficient approaches to solve the visibility problem have already been developed in computer graphics. There occlusion detection is required within the virtual images during scene rendering. These approaches can be modified for efficient texture extraction and placement in the context of cultural heritage representation. As an example, the algorithm proposed by (Grammatikopoulos L. et al 2005) is based on z-buffering, which computes the visibility of each pixel by comparing the surface depths at each position in image space and selects the one closer to the observer. (Alshawabkeh & Haala 2005) combine such an approach with techniques originating from the painter's algorithm. In this approach, also well known in computer graphics for occlusion detection, the model polygons are sorted by their depth in object space with respect to the collected texture image. This algorithm efficiently detects ambient, back-face and view frustum occlusions based on computations both in image and object space.

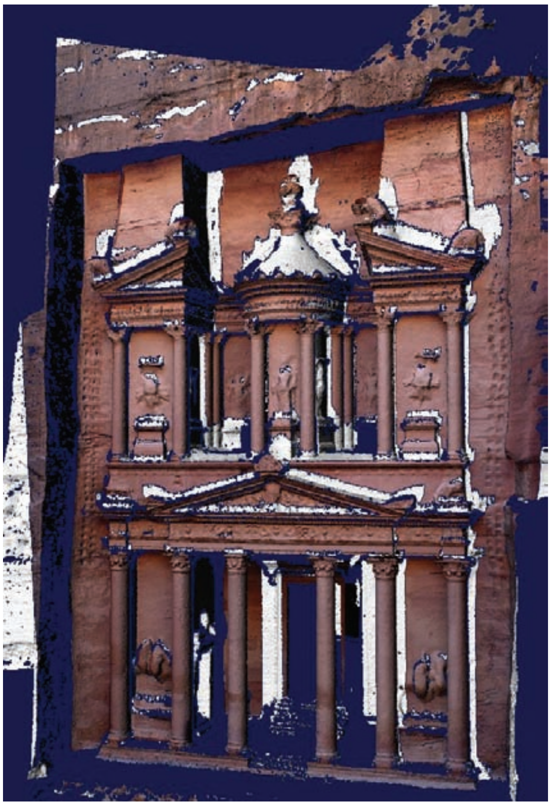


Figure 5: *Result of texture mapping with occlusions detection.*

An intermediate result of this texture mapping process for the meshed 3D model of the Alkasneh monument is depicted in Figure 5**.** In this example, the image already shown in Figure 4d is used to provide the required surface texture. In Figure 5 the parts of the 3D model, which are not visible in this texture image are marked as grey regions. These regions are separated from the original meshed 3D model and then used as input surfaces for a second texture image. The process for visibility checking and model separation is repeated until the 3D model is textured using all available images. Finally, the different parts are merged again to provide the overall model.

This process is demonstrated additionally in Figure 6 to Figure 9 for a data set showing a Roman Theatre in Jerash, Jordan. For this project, laser data was collected from 6 stations by a Mensi GS100 scanner, resulting in a 3D model consisting of 4.6 million triangles. Figure 6 shows the digital image, which was used for the texture mapping depicted in Figure 7. Similar to Figure 5**,** the parts of the 3D model, which are not visible, are marked in grey.

Figure 6: First texture image for North theatre –Jerash

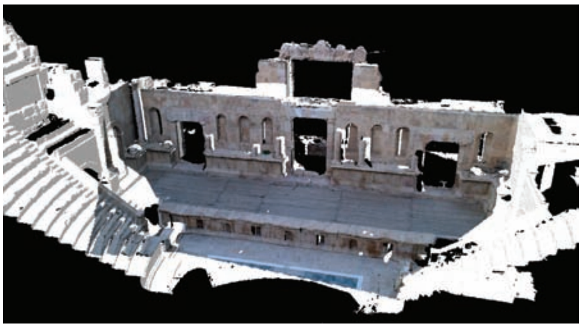


Figure 7: 3D model textured by first image after occlusion detection

As it is shown in Figure 8 these parts are then extracted from the original model and used for further processing. In Figure 8, the image depicted in Figure 9 has already been used for this purpose. After this step, those meshes of the 3D model, which are not visible in the image are again extracted and have to be textured from another available image in subsequent steps. Finally, the separated parts are merged again to generate the overall model depicted in Figure 11**.**

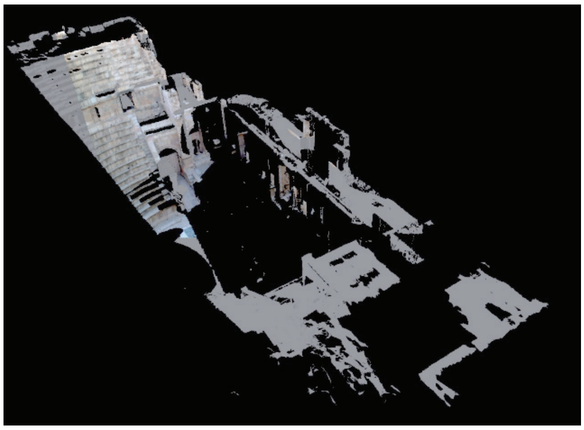


Figure 8: Occluded parts in first image as extracted from the 3D model after texture mapping from second image.

Figure 9: *Second texture image.*

Figure 10: *Additional parts of textured 3D model.*

Additional parts of the 3D model, which have been textured and separated by different images, are depicted in Figure 10. In addition to the correct geometric processing of the images, which is realised by the co-registration process discussed in section 2, homogenous radiometry between the different texture images has to be guaranteed to allow for realistic visualisation of the collected 3D model. Optionally, artefacts due to illumination changes have to be removed in an additional pre-processing step. In our application an off-the-shelf remote sensing image processing package was used for histogram equalization and stitching of the texture images. The final texture model is depicted in Figure 11.

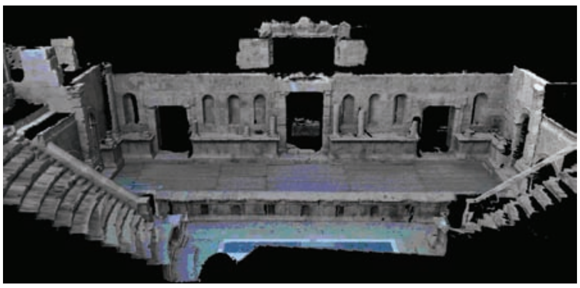


Figure 11: *3D textured model using 4 images after colour correction.*

4. Conclusion

Especially for the documentation of complex terrestrial scenes, the 3D geometric model of real world objects has to be enhanced with texture as provided from separate sets of photographs. Thus, the generation of 3D virtual models in the context of cultural heritage documentation frequently requires the combination of terrestrial LIDAR and image data in order to optimize their geometric accuracy and the visual quality.

The combination of the different data sets has to be realized by an exact co-registration, which is implies a 2D-3D pose estimation algorithm. The most common methods for solving such registration problems between two datasets are based on the identification of corresponding points. However, such methods usually are not applicable when dealing with surfaces derived from 3D point clouds. This type of data is derived from laser footprints with limited scanning resolution rather than distinct points that could be identified in the imagery. Thus, within point clouds and range images the perception of objects structure is limited and not very appropriate for registration.

To allow for a precise and reliable co-registration of the data sets, linear features were extracted by a suitable segmentation process. Since these lines possess a considerable amount of semantic information, the correspondence problem between the image and object space can then be solved easier. Currently, the segmented lines are used as input for a following manual selection of correspondences, while future work will aim on an automatic matching process.

The final goal of our work is the generation of photo-realistic models of complex shaped heritage sites with optimal efficiency. In addition, to the co-registration process, this is supported by a fast algorithm to verify the visibility of the available texture images. While these algorithms are currently used in a post-processing mode, an easy visibility detection of 3D models and 2D images should also be beneficial to guarantee the completeness and sufficient coverage during data collection.

References

[AH05] ALSHAWABKEH, Y. & HAALA N.: Automatic Multi-Image Photo Texturing of Complex 3D Scenes. *CIPA IAPRS Vol. 34-5/C34* (2005), pp. 68-73.

[BJ88] BESL, P. & JAIN, R. Segmentation Through Variable-Order Surface Fitting: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(2), (1998) pp. 167–192.

[CH96] CHRISTY, S. & HORAUD, R.: Iterative Pose Computation from Line Correspondences. *Computer Vision and Image Understanding* Vol. 73, No. 1, January, (1996) pp. 137–144.

[GKK*05] GRAMMATIKOPOULOS L., KALISPERAKIS I., KARRAS G. & PETSA E.: Data Fusion from Multiple Sources for the production of Orthographic and Perspective Views with Automatic Visibility Checking. *CIPA IAPRS Vol. 34-5/C34* (2005)

[HJJ*96], HOOVER, A., JEAN-BAPTISTE, G., JIANG, X, FLYNN,P. & BUNKE, H.: An Experimental Comparison of Range Image Segmentation Algorithms. IEEE *Transactions on Pattern Analysis and Machine Intelligence,* 18(7) (1996) pp. 673-689.

[KF03] KLINEC, D. & FRITSCH, D.: Towards pedestrian navigation and orientation. *Proceedings of the 7th South East Asian Survey Congress: SEASC'03*, Hong Kong, (2003).

[LH02] LEE, Y. & HABIB, A.: Pose Estimation of Line Cameras Using Linear Features, *ISPRS Symposium of PCV'02 Photogrammetric Computer Vision,* Graz, Austria (2002).

[LHS00] LENSCH, H., HEIDRICH, W. & SEIDEL, H.: Automated texture registration and stitching for real world models. In *Proceedings of Pacific Graphics* (2000), pp. 317–326.

[LS05] LIU, L. & STAMOS, S.: Automatic 3D to 2D registration for the photorealistic rendering of urban scenes. *Computer Vision and Pattern Recognition. IEEE Computer Society.* (2005)

[MLM01] MARSHALL, D., LUKACS, G. & MARTIN, R.: Robust segmentation of primitives from range data in the presence of geometric degeneracy. *IEEE Trans. Pattern Analysis and Machine Intelligence* 23(3), (2001), pp. 304-314.

[SD01] SAPPA, A. & DEVY, M.: Fast Range Image Segmentation by an Edge Detection Strategy: *Proc. IEEE Conf. 3D Digital Imaging and Modelling*, (2001), pp. 292-299.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Automatic Registration and Calibration for Efficient Surface Light Field Acquisition

Frédéric Larue[†] and Jean-Michel Dischler[†]

LSIIT UMR CNRS-ULP 7005, Université Louis Pasteur, Strasbourg I, France

Abstract

This paper presents a novel protocol for the acquisition of surface light fields which is designed to deal with delicate objects that might not be touched or moved. This constraint is particularly important when art pieces are involved. Our protocol enables the automatic reconstruction of a model from many range images and the automatic registration of many pictures with the acquired geometry. A structured light pattern is first used to project a parameterization over the analyzed surface. Each surface point hit by this parameterization is uniquely identified, independently of the chosen viewpoint, and the problem of finding point-point and point-pixel correspondences is then immediately solved. These correspondences are finally used to perform the registrations and camera calibrations that provide the data to be used by a surface light field renderer.

1. Introduction

In the research field of computer graphics and visualization, a part of the scientific community is attempting, for many years, to take account of reality to increase the visual quality of synthetic images. The appearance and the development of digitization tools have widely promoted this kind of approach, allowing numerical measurements of complex real data. Unfortunately, although these tools are greatly used nowadays, they are often subject to high constraints and such measurements are not always easy.

The task is even more difficult when additional constraints arise from the objects to be measured. This is the case of our work which is a part of a national project done in conjunction with the ministry of culture and with museums. One aim of this project is the archiving of art pieces by the establishment of a numerical imprint, including geometrical and photometrical information. We are then interested in capturing the shape and the appearance of fragile models that might not be touched or moved too often.

Concerning the geometry, current devices are not able to immediately acquire the whole surface of an object. Range scanners, for example, can only view one side of the object at a time, and a complete digitization requires several acquisitions by placing the scanner at different locations to cover its surface as best as possible. All measured surface parts must be post processed during a reconstruction step. As each scan is defined in the scanner local frame, the first problem addressed by reconstruction is to express all scans in a common global frame. This problem, called *registration*, can be

Figure 1: Left*: a picture of the Greek vase model.* Middle*: a model reconstructed from several range images.* Right*: a synthetic view generated from the surface light field captured with our method. All acquired range images and pictures are registered in a fully automatic manner.*

[†] e-mail: {larue, dischler}@lsiit.u-strasbg.fr

easily solved by systems such as robotic arms, able to accurately localize the scanner with respect to its target. All motions are recorded, and the transformation associated to each scan is immediately known. In spite of their convenience, such devices are expensive or are not always designed to be displaced. If the motion of the scanner with respect to the object is not recorded, it must be estimated. The crucial point to solve for an efficient registration solution is then to be able to accurately determine some correspondences between geometric points of the different scans.

Another important part of the digitization of real surfaces is concerned with the acquisition of materials. Indeed, as illustrated in figure 1, the only shape is not sufficient to represent the digital copy of an object in a realistic manner. Emerging in parallel to digitization tools, dedicated rendering techniques have been developed. Among these, the surface light fields attempt to represent the appearance of an object within a fixed lighting environment and from an arbitrary viewpoint. In order to synthesize images from real data, the radiance emitted by the considered object has to be measured beforehand for many viewpoints. This information is commonly captured by taking several pictures from different viewpoints. To correctly interpret this captured appearance, the viewing direction associated to each picture must be known. It can be determined by solving a well known problem of camera calibration, where intrinsic (optical parameters) and extrinsic (camera pose) parameters are estimated. As this estimation is computed from the image space projections of many scene points, the efficiency of the calibration procedure, once again, depends on the ability to accurately establish some correspondences between two data sets: the geometric points and their matching camera pixels.

This paper proposes a new protocol for the acquisition of geometry and radiance specifically designed to deal with delicate models. Neither contact nor displacement of the measured object is involved and all the registration procedures are fully automatic. Concerning the remainder of the paper, the related scientific context is first explored in section 2. An overview is presented in section 3 and the technical points are then explained in sections 4 and 5, respectively describing our extraction of correspondences and its use for the acquisition of surface light fields. Results and studies are presented in section 6, followed by conclusions in section 7.

2. Related Work

A light field is an approximation of the plenoptic function [AB91] which describes for all points in space the incident light incoming from the whole scene. The first approaches proposed to represent this function were purely image based renderings [LH96,GGSC96], able to generate new views from a set of acquired pictures. Later, the *surface light fields* [MRP98, CBCG02] propose to store the light field directly over the surface of an object, leading to some interesting simplifications. In first, only the relevant information is kept, avoiding all background data. Next, as the information is stored on the surface, texture mapping can be exploited to speedup the rendering. Unfortunately, as a surface light field requires a geometric support, both the radiance and the shape must be acquired while dealing with real data. Hence, the registration problem is of great importance and is present at two different levels: many range images must be merged to produce a single consistent model and many pictures which sample the radiance must be registered with the acquired geometry to determine the data of the surface light field.

Concerning the geometry, many works have investigated the problem of reconstructing a single model from many range data. When an initial coarse alignment is known, iterative methods are able to progressively refine the solution [BM92,TL94,BS99,GGT00,GG01]. But if a fully automatic procedure is preferred, such a prior knowledge is not always available and some correspondences must then be found to compute a transformation between different data sets. Based on the idea that the scanner pose is never arbitrarily chosen, a knowledge about the adopted scanning strategy [PFC*05] enables to predict the overlaping relationships between scans, reducing the search of matching elements to small subsets. Sometimes, features may be extracted when data arises from specific scenes or situations. Urban scenes [ZSHQ04], for examples, present many apparent and organized edges that can be identified. In the case of a real time acquisition pipeline [RHHL02], the temporal coherence between successive frames can be exploited. But even if feature extraction has the advantage of avoiding the requirement of a prior knowledge, it is generally designed for scene dedicated methods and not for general approaches. Based on the same idea, other works attempt to extract invariant characteristics which does not rely on any assumption about the scene [JH97, CHC98, ZH99]. These methods, even if working on arbitrary data sets, are often based on exhaustive searches and are then computationnaly expensive. All the aforementioned techniques are focused on pair-wise registration, only able to deal with two data sets. Generally, a complete digitization requires many more than two acquisitions and global registration methods have been proposed to take into account all of the resulting range images at the same time [Pul99,HH01,NI02,ZSHQ04]. Unfortunately, many of them require an initial alignment to prove practicable or efficient.

Beyond the shape, a surface light field acquisition protocol has to be able to register a set of pictures with the geometry. The most common way to achieve this is to use targets [CBCG02]. The problem is that targets must be seen by the acquisition devices and this is not an obvious task. In cultural heritage, for example, scanning art pieces forbids to put targets directly over the object. In its neighborhood, some occlusion problems may be introduced. Moreover, depending on lighting conditions, an automatic image segmentation may fail to localize the targets. Methods based on silhouette matching [MK99] might not be robust enough

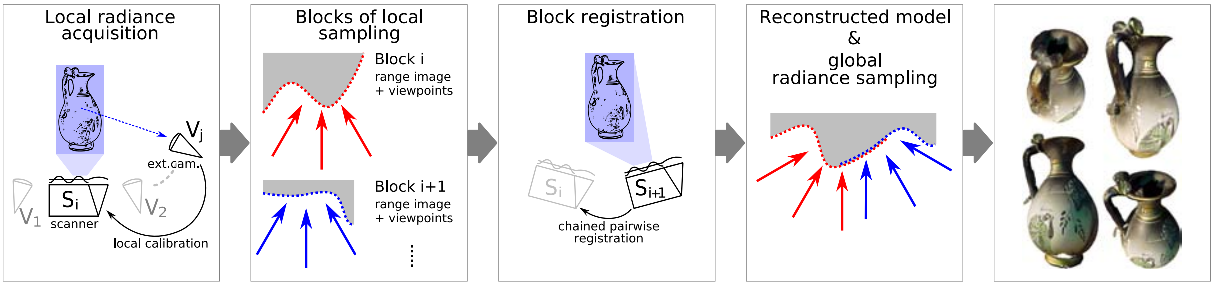


Figure 2: *Our acquisition protocol. A local radiance sampling is first acquired by registering several pictures with respect to the current range image. The resulting* blocks of local sampling *are then merged together in a common global frame by a chained procedure that register each block with respect to the previous one. The registration transformations are applied to the range images and to their associated sets of pictures, leading to a consistent model and a global radiance sampling.*

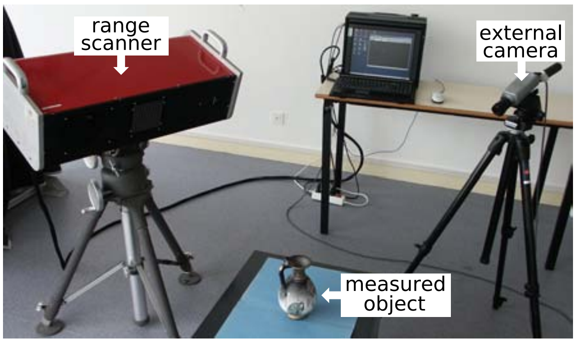


Figure 3: *Our digitization bench to capture whole surface light fields. Only a lightweight device is involved: a structured light range scanner and an external camera.*

(with surfaces of revolution or symmetrical objects, for example). Most recently, a system has been proposed to infer new image-to-geometry correspondences from a set of known ones [FDG*05]. But the user interventions, even if greatly reduced, are not totally avoided as the initial set of correspondences must be specified manually.

3. Method Overview

We are interested in capturing the appearance of delicate objects, such as art pieces, that cannot be touched or moved. At the end of the acquisition step, we recover all the data required by a surface light field rendering method, that is a fully reconstructed model and a set of pictures, to sample the radiance, whose viewpoints are known. To achieve this goal, we propose an acquisition protocol which performs the model reconstruction and the viewpoint determination in a fully automatic manner without any contact or displacement of the measured object. Moreover, to agree with a mobility constraint, only a lightweight hardware is involved: we just need a range scanner based on structured light and an external digital camera, as shown in figure 3.

Our protocol, summarized in figure 2, works as follows: an acquisition procedure enables to automatically register many pictures with respect to a single range image. This step is iterated as mush as needed to cover all the object surface, resulting in many separate blocks made of a range image and its associated set of pictures. All blocks are then registered by a chained procedure that register each new acquired block with respect to the previous one. The transformations required to align the scans are automatically computed and are applied not only to the range images but also to the associated sets of calibrated pictures. Thus, both the geometry and the pictures used to sample the radiance are expressed in a common global frame.

The two tasks of registering a picture with a known geometry and registering a piece of surface with another one consist in computing a transformation between two data sets from a list of correspondences which must be determined. The major benefit of this work is the solution proposed to solve this relationship problem. We use a structured light pattern to project a 2D parameterization onto the considered object. Consequently, all the surface points covered by this parameterization are identified by a unique couple of coordinates. The search of correspondences then reduces to find points in the different data sets whose parameterization coordinates are equal. The picture's viewpoints and the range image's alignments are then computed from these correspondences by some well known and experienced algorithms (see sections 5.1 and 5.2).

As our goal is the acquisition and not the visualization, we are using a basic surface light field renderer to provide some examples to rely on for analysis and validation.

4. Extraction of Correspondences

In this section, we first discuss the interesting properties of the structured light model we use and how we extend it

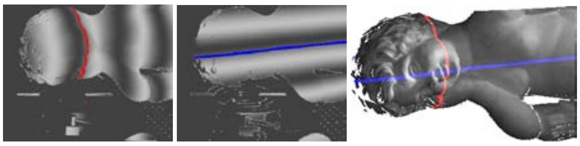


Figure 4: *The phase is strictly increasing orthogonally to the stripes orientation, producing some iso-phase lines over the measured surface. Given two different stripes orientations, each surface point is associated to the intersection of two iso-phase lines and is then uniquely identified.*

to produce a spatial parameterization of the measured surface. This parameterization defines a unique identifier at each point and is used to establish a set of correspondences between different data sets. As some errors are necessarily introduced while measuring, some of the correspondences which have been found are erroneous and it is of great importance to estimate the accuracy of the acquired data. We therefore introduce an error metric based on our parameterization to easily and efficiently classify the possible outliers.

4.1. Structured light model properties

The structured light model we use is based on the phase shifting principle. A set of grayscale stripes whose intensity's variation follows a sinusoidal is projected and an image analysis enables to determine, for a pixel p of a sensor camera, the value of the phase $\phi(p)$ corresponding to the observed surface point. Many shape measurement methods, as [HZ05] for example, are based on a phase shifting principle to compute a depth information by optical triangulation. We recommend the reader to take a look at these works for more precise informations.

An interesting property is that such structured light models induce a 1D parameterization of the measured surface. Indeed, the phase information is monotonically increasing and is continuously defined over the whole projection domain, orthogonally to the stripe direction. Each stripe is then clearly identified by a unique phase value, producing some iso-phase lines over the object. A second property is that the computed phase is independent of the viewpoint chosen to capture it. If the projector remains static with respect to the measured surface, the phase values computed at a given surface point from different viewpoints are identical, whatever the camera settings are.

4.2. Extension to a surface parameterization

The 1D parameterization induced by this structured light model can be easily extended to a 2D parameterization. By projecting the same stripe pattern along two orthogonal orientations, a couple $\Phi(p) = (\phi_1(p), \phi_2(p))$ of phase values is defined at each surface point p. As the two functions $\phi_1(p)$ and $\phi_2(p)$ are monotonically increasing over their own domains and then do not have the same value twice, the couple $\Phi(p)$ represents a unique identifier for the point p, as illustrated in figure 4. In the remainder of this paper, we call this 2D parameterization the Φ-*parameterization*.

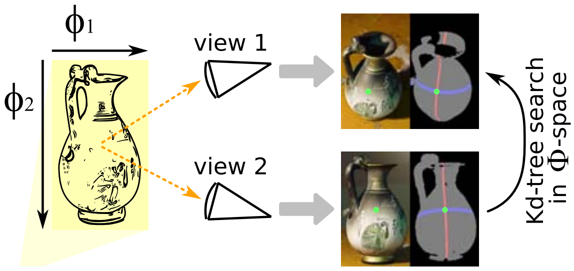


Figure 5: *The same Φ-parameterization is captured (the projector and the object remain static) from two different viewpoints to define correspondences. An element in the first view is paired with the element whose couple of phases is the nearest in the second view.*

By considering the properties of the structured light model presented beforehand, as long as the scanner and the considered object remain static, the Φ-parameterization remains the same and is completely independent of the viewpoint chosen for its acquisition. As a consequence, while the parameterization does not change, two pixels p and q coming from two distinct viewpoints and having their coordinates $\Phi(p)$ and $\Phi(q)$ identical are necessarily focused on the same surface point.

4.3. Selection of correspondences

This identification of some surface points is used in order to solve the problem of finding correspondences between different data sets. Unfortunately, practical problems arise as digitization tools are subject to many error sources ([RHHL02]). The most stringent one is the CCD discretization: as the acquisition camera is not able to capture a continuous domain, two pixels taken from different viewpoints never see exactly the same surface region, involving dissimilarities between their respective Φ coordinates. The search of correspondences is then no longer an equality test but should be replaced by a nearest neighbor search.

We use the squared Euclidean distance between the couples of phases as an accuracy criterion. Given two points x and y coming from different viewpoints but captured with the same Φ-parameterization and their respective couples of phases $\Phi(x) = (\phi_1(x), \phi_2(x))$ and $\Phi(y) = (\phi_1(y), \phi_2(y))$, the distance between x and y is denoted $\varepsilon(x,y)$ and is defined by equation 1:

$$\varepsilon(x,y) = (\phi_1(x) - \phi_1(y))^2 + (\phi_2(x) - \phi_2(y))^2 \quad (1)$$

As shown in figure 5, given two data sets identified by

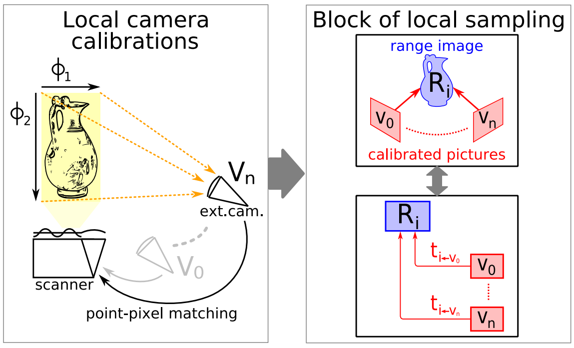


Figure 6: *The acquisition of a* block of local sampling. *The block is made of the range image and all the pictures that have been calibrated with respect to it.*

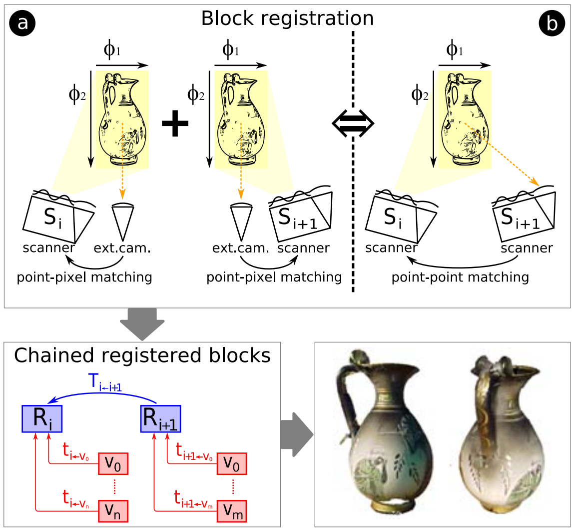


Figure 7: *A local sampling block is registered by a chained procedure with respect to the previously acquired one. The external camera is used as a fixed reference between the two scanner positions.*

the same Φ-parameterization, the correspondences are then found by parsing all elements of the first set and searching the nearest element with respect to the ε-distance in the second one. This search is efficiently implemented by using Kd-trees. As we are exploring the space of the Φ-parameterization, trees of dimension two are used.

4.4. Outlier classification

Among all the resulting pairs, some might not be valid. Indeed, even if the Kd-tree search leads to a result, the nearest element that has been found is not necessarily a good correspondence. We use the ε-distance to determine the validity of each association. If the ε-distance exceeds a given threshold S_ε, the two elements are considered as too distant and the association is discarded. More than rejecting outliers, this thresholding is a good way to retain only the most accurate correspondences by setting a low threshold. In our application, this threshold is data dependent and is defined as $S_\varepsilon = \lambda\, S$, where S is the average ε-distance between adjacent pixels of the considered viewpoint and λ is a factor depending on the desired quality of registration.

5. Surface Light Field Acquisition

As our acquisition protocol is based on the aforementioned principle, we have modified our range scanner to be able to get back the phase image before its transformation into a range image. A second modification enables to perform a phase acquisition with an external camera instead of only the one embedded in the scanner. The protocol is decomposed in two parts. In first, a local sampling of the radiance consisting in a set of calibrated pictures is linked to each acquired range image, resulting in what we call *local sampling blocks*. Next, all these blocks are merged together by a registration step that remaps all data (geometry and pictures) in a common global frame.

5.1. Acquisition of a local sampling block

Considering a single range image R_i, the acquisition of the associated radiance information is easily made by using our structured light pattern, as illustrated by figure 6. After having performed the geometric acquisition, the scanner projects the Φ-parameterization onto the object and captures it. As the scanner as not been moved, there is a pixel-to-pixel matching between the phase image and the range image R_i, and many acquired surface points can then be uniquely identified, as explained in section 4.2. The radiance information is then captured by using the external camera. More than taking the picture v_j, this camera also captures the Φ-parameterization. Some correspondences can then be established between many pixels of v_j and the matching 3D points of R_i by using the search procedure of section 4.3. Tsai's calibration algorithm [Tsa92] finally estimates the scan-to-camera transformation $t_{i \leftarrow v_j}$ from these point-pixel correspondences and thus associates a viewing direction and a localization to the considered picture v_j.

This procedure is repeated for many camera positions to get a radiance information with a dense sampling of viewing directions. Obviously, the external camera needs to view the parameterization projected from the current scanner position. Thus, the captured set of pictures does not correspond to a whole radiance sampling but only to a part of it, defined around the current scanner viewpoint. We call *local sampling block* the set composed of the range image R_i and all the pictures $v_0, \ldots, v_n$ that are locally calibrated with respect to it.

5.2. Block registration

Each block represents an isolated part of the final surface light field, disconnected from the others. All blocks must then be merged together by a registration process. This registration does more than just aligning the geometry data sets: a rigid transformation is computed from the geometric data and is applied to the whole local sampling block, that is to say to the range image and to its set of locally calibrated pictures. All geometry parts are then merged in a common global frame, as well as all calibrated pictures.

Each time a new block is to be acquired, the scanner is moved to a new position. However, there is no explicit common reference between the frames of the different blocks. We are then using the external camera as a fixed reference between two successive block acquisitions, as illustrated by figure 7. Standing at a given position, the external camera captures the two parameterizations projected from the previous and the current scanner positions. As previously discussed, correspondences are extracted between some pixels of the external camera and some points of both range images R_i and R_{i+1}. The pixels that are linked to the both range images provide the geometric correspondences needed for the block registration. The rigid transformation $T_{i \leftarrow i+1}$ to remap R_{i+1} in the frame of R_i is finally computed from these correspondences using a quaternion-based method [BM92]. Once the registration transformation is known, we apply it not only to the range image but also to the viewpoints of the associated set of pictures. For the i-th block composed of the range image R_i and the viewpoints $v_0, \ldots, v_n$, the transformation T_i to remap R_i to the global frame is the composition of all the previous registration transformations $T_i = T_{1 \leftarrow 2} \times \ldots \times T_{i-1 \leftarrow i}$ and the transformation t_j that projects the global frame to the image space of the viewpoint v_j is defined as $t_j = t_{i \leftarrow v_j} \times T_i^{-1}$. Thus, each block is registered with respect to the previous one by a chained procedure. The final geometry reconstruction is performed by the VRIP algorithm [CL96].

If two range scanners are available, a more immediate solution is clearly possible. The second scanner can act as the external camera: it captures the parameterization projected by the first scanner, as shown is figure 7*b*. The phase based search then results in a direct mapping between the two scanner frames. It is obvious that using an intermediate device instead of two scanners clearly induces a loss of accuracy. This point is discussed in the result section. Nevertheless, it is important to note that the examples provided in this paper have been produced with the method involving only one scanner. Usage and measurements have shown that this first solution is accurate enough to be used.

6. Results

To visualize the data provided by our method, we have developed a basic rendering algorithm which computes the color of a geometric primitive for a given viewing direction by the interpolation of the three closest radiance samples. Figures 1 and 8 have been generated by this renderer. These surface light fields have been reconstructed from 5 range images and 27 viewpoints for the Greek vase model, and from 6 range images and 42 viewpoints for the African wood statue.

Concerning the calibration process, the main advantage of our method against the use of standard targets resides in the number of available calibration points. While the number of targets in the scene is necessarily limited, the number of available point-pixel correspondences is generally not exceeding a few tens. In our application, the number of calibration points used in the Tsai's algorithm is of many thousands. Moreover, as we use a projected parameterization, the occlusion problems that can be encountered with targets are avoided. Concerning the registration, table 1 compares ICP to our method in terms of accuracy. As can be seen, ICP remains more accurate, certainly due to its iterative nature. But it is important to recall that ICP requires an initial coarse alignment whereas our registration if fully automatic. Moreover, ICP may fall into a local minimum if the two surfaces

Figure 8: Left*: the mesh of the Venus at Bath reconstructed from 23 range images registered with our method.* Right*: a picture of the African statue and two synthetic views generated from the surface light field acquired with our protocol.*

Set	ICP		Φ-param.	
	mean	*std. dev.*	*mean*	*std. dev.*
Angel	0.270	0.235	0.328	0.235
Greek1	0.234	0.360	0.292	0.371
Greek2	–	–	0.234	0.336
African	0.248	0.265	0.250	0.262

Table 1: *Comparison of ICP against our method. The given values correspond to the average distance (in mm) between nearest neighbors in the overlapping region of the two scans. Empty cells correspond to a case where ICP has failed to perform the registration.*

ICP		Intermed. camera		Two scanners	
mean	*std. dev.*	*mean*	*std. dev.*	*mean*	*std. dev.*
0.270	0.235	0.328	0.235	0.303	0.239

Table 2: *Comparison of our two variants of registration. The given values correspond to the average distance (in mm) between nearest neighbors in the overlapping of two scans.*

Nb. points in the 1st scan	*Nb. points in the 2nd scan*	*Nb. corres. found*	*Registration time (ms)*
325K	331K	15K	629
331K	329K	2K	455
75K	76K	3K	419
215K	182K	10K	579
23K	20K	11K	250

Table 3: *Timings measured during some pairwise registrations. The registration time includes the search of correspondences and the computation of the rigid transformation.*

present the same global shape. We have experienced this problem with the Greek vase model, as shown in table 1. We have also compared the two variants of our protocol (with one scanner and a camera or with two scanners). Two scanners have performed a geometric acquisition and an external camera were placed between them during the phase acquisitions. The result, reported in table 2, shows a loss of accuracy induced by the use of the external camera instead of the two scanners. This loss was predictable but is not as significant as we expected.

In terms of performance, registration timings are given in table 3. These timings have been obtained with a processor AMD Athlon 3800+. The bulk of our technique consists in finding inside a range image the best approximation of a given point, based on its phase identifier. This search is done in the space of the Φ-parameterization and must be done only once, as opposed to the ICP algorithm where Kd-trees of dimension three must be recomputed for each iteration. Performance for the calibration has not been measured, as it only depends on the effectiveness of the Tsai's algorithm.

There are two drawbacks with our method. The main one is the cumulative nature of the error due to the chained pairwise registration. We have compared, in table 4, the average distance between the two scans of all the registered pairs and the distance between the first and the last scans of the whole chain of 23 range images of the Venus at Bath. In this example, the incidence of the accumulation remains neglectable as it does not induce any misalignment artefact during the reconstruction. This first drawback exists in all methods that are not designed in the purpose of a global registration. However, our results can be used as a good starting point for global registration methods where an initial alignment is needed. As shown in table 3, our registration is fast enough to be used as the initialization of another technique.

Avg. dist. for pairs: 0.243mm	*Last-to-first dist.:* 0.477mm

Table 4: *Error accumulation for the 23 acquisitions of the Venus at Bath.* Left*: average distance of all registered pairs.* Right: *distance between the first and the last range images.*

The second drawback is related to the capture of the radiance. Since the Φ-parameterization must be known for each viewpoint, the method does not allow the use of a hand held camera. Indeed, both the camera and the projector have to remain static since a picture and two phase acquisitions need to be taken with respect to the exact same viewpoint. The acquisition time of the radiance information may then be increased compared to the use of standard targets. Moreover, as the calibration process depends on the registration to establish a global radiance sampling, the cumulative error described beforehand may have an incidence on the accuracy of the viewpoint localization.

7. Conclusions & Future Work

We have presented a new protocol for the acquisition of surface light fields from real objects. This protocol is designed to perform some measurements on delicate objects, such as art pieces in a context of cultural heritage, that cannot be touched or moved. We are using a structured light pattern to project a parameterization over the analyzed surface which enables to uniquely identify many scene points. This identification is used to deduce the viewing directions of a set of pictures that captures the radiance of the scene, but also to perform a chained pairwise registration to reconstruct a consistent model from many range images.

Due to the use of a spatial parameterization, a quasi immediate mapping is established between the different data sets. The search of correspondences and so the registration are then fast comparing to iterative methods, even if the result is not as accurate. To increase accuracy, our solution could be used as a good starting point for a global registration method to avoid the cumulative nature of the error.

Concerning the material acquisition, we work on extending our protocol to enable the digitization of models with their complete bidirectional information. We are particularly interested in the simplification of the digitization process as it is always, nowadays, a really tedious task due to the requirement of an exhaustive sampling of the lighting. Indeed, capturing a bidirectional information means to control the lighting environment. The ability to precisely localize a light source in order to deduce the incident light directions is then another goal to achieve.

8. Acknowledgement

This work is funded by the french ministry of research (*RIAM*-project *AMI3D*, no. 04 C 292).

References

[AB91] ADELSON E., BERGEN J.: The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing* (1991).

[BM92] BESL P. J., MCKAY N. D.: A method for registration of 3-d shapes. *IEEE Trans. Pattern Anal. Mach. Intell. 14*, 2 (1992), 239–256.

[BS99] BENJEMAA R., SCHMITT F.: Fast global registration of 3d sampled surface using a multi-z-buffer technique. In *Image and Vision Computing 17* (1999), pp. 113–123.

[CBCG02] CHEN W.-C., BOUGUET J.-Y., CHU M., GRZESZCZUK R.: Light field mapping: efficient representation and hardware rendering of surface light fields. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (2002), pp. 447–456.

[CHC98] CHEN C.-S., HUNG Y.-P., CHENG J.-B.: A fast automatic method for registration of partially-overlapping range images. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision* (1998), p. 242.

[CL96] CURLESS B., LEVOY M.: A volumetric method for building complex models from range images. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (1996), pp. 303–312.

[FDG*05] FRANKEN T., DELLEPIANE M., GANOVELLI F., CIGNONI P., MONTANI C., SCOPIGNO R.: Minimizing user intervention in registering 2d images to 3d models. *The Visual Computer 21*, 8-10 (sep 2005), 619–628. Special Issues for Pacific Graphics 2005.

[GG01] GREENSPAN M., GODIN G.: A nearest neighbor method for efficient icp. In *3DIM'01: Third International Conference on 3-D Digital Imaging and Modeling* (2001), pp. 161–168.

[GGSC96] GORTLER S., GRZESZCZUK R., SZELISKI R., COHEN M.: The lumigraph. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (1996), pp. 43–54.

[GGT00] GREENSPAN M., GODIN G., , TALBOT J.: Acceleration of binning nearest neighbor methods. In *Vision Interface 2000* (2000).

[HH01] HUBER D., HEBERT M.: Fully automatic registration of multiple 3d data sets. In *IEEE Computer Society Workshop on Computer Vision Beyond the Visible Spectrum (CVBVS 2001)* (Dec 2001).

[HZ05] HUANG P. S., ZHANG S.: A fast three-step phase-shifting algorithm. In *Proceedings of SPIE, OpticsEast* (2005).

[JH97] JOHNSON A. E., HEBERT M.: Surface registration by matching oriented points. In *NRC '97: Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling* (1997), p. 121.

[LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (1996), pp. 31–42.

[MK99] MATSUSHITA K., KANEKO T.: Efficient and handy texture mapping on 3D surfaces. In *Computer Graphics Forum (Eurographics '99)* (1999), vol. 18(3), pp. 349–358.

[MRP98] MILLER G., RUBIN S., PONCELEON D.: Lazy decompression of surface light fields for precomputed global illumination. In *Rendering Techniques '98* (1998), pp. 281–292. Proc. 9th Eurographics Rendering Workshop, Vienna, Austria, June 29 – July 1, 1998.

[NI02] NISHINO K., IKEUCHI K.: Robust simultaneous registration of multiple range images. In *Proceedings of Fifth Asian Conference on Computer Vision ACCV'02* (2002), pp. 454–461.

[PFC*05] PINGI P., FASANO A., CIGNONI P., MONTANI C., SCOPIGNO R.: Exploiting the scanning sequence for automatic registration of large sets of range maps. *Computer Graphics Forum 24*, 3 (2005), 517–526.

[Pul99] PULLI K.: Multiview registration for large data sets. In *International Conference on 3-D Digital Imaging and Modeling* (1999), pp. 160–168.

[RHHL02] RUSINKIEWICZ S., HALL-HOLT O., LEVOY M.: Real-time 3d model acquisition. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (2002), pp. 438–446.

[TL94] TURK G., LEVOY M.: Zippered polygon meshes from range images. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (1994), pp. 311–318.

[Tsa92] TSAI R. Y.: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. In *Radiometry* (1992), pp. 221–244.

[ZH99] ZHANG D., HEBERT M.: Harmonic maps and their applications in surface matching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '99)* (1999), vol. 2.

[ZSHQ04] ZHANG A., SUN W., HU S., QIAN C.: Automatic global registration of multiple 3d data sets from outdoor urban environments based on feature units. In *SCCG '04: Proceedings of the 20th spring conference on Computer graphics* (2004), pp. 193–199.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

High Quality PTM Acquisition: Reflection Transformation Imaging for Large Objects

Matteo Dellepiane, Massimiliano Corsini, Marco Callieri, Roberto Scopigno

Visual Computing Lab, ISTI - Area della ricerca CNR
Via G. Moruzzi 1, 56124, Pisa, Italy
Email: {dellepiane, corsini, callieri, scopigno}@isti.cnr.it

Abstract

Reflection Transformation Imaging has proved to be a powerful method to acquire and represent the 3D reflectance properties of an object, displaying them as a 2D image. Recently, Polynomial Texture Maps (PTM), which are relightable images created from a set of photos of the object taken under several different lighting conditions, have been used in Cultural Heritage field to document and virtually inspect several sets of small objects, such as cuneiform tablets and coins. In this paper we explore the possibility of producing high quality PTM of medium or large size objects. The aim is to analyze the acquisition pipeline, resolving all the issues related to the size of the object, and the conditions of acquisition. We will discuss issues regarding acquisition planning and data gathering. We also present a new tool to interactively browse high resolution PTMs. Moreover, we perform some quality assessment considerations, in order to study the degradation of quality of the PTMs respect to the number and position of lights used to acquire the PTM. The results of our acquisition system are presented with some examples of PTMs of large artifacts like a sarcophagus of 2.4 × 1 m size. PTM can be a good alternative to 3D scanning for capturing and representing certain class of objects, like bas-relieves, having lower costs in terms of acquisition equipment and data processing time.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation I.3.4 [Computer Graphics]: Graphics Utilities I.3.6 [Computer Graphics]: Methodology and Techniques

1. Introduction

Reflection Transformation Imaging has proved to be a very interesting and powerful method to acquire and represent the 3D reflectance properties of an object, displaying them on a 2D image. Typically, this is done by calculating an approximation of the reflectance function of the objects' surface for each pixel of the final image. This approximation is calculated starting from a set of photos of the object, each one taken under controlled illumination. One of the most popular techniques for reflection transformation imaging is Polynomial Texture Mapping (PTM) [MGW01]. Several applications of this technique have been proposed, mainly in the field of Cultural Heritage. The use of contrast enhancement mechanisms (specular enhancement, diffuse gain) proved to be very useful not only in term of documentation, but also in terms of analysis of the surface features. Several features which were not visible during physical inspection were discovered by this representation. Nevertheless, until now the method has been applied only to small sized objects. In this paper we explore the possibility of producing high quality PTMs of medium-large objects, from 60-70 cm of width by 50-60 cm of height up to 2m of width by 1m of height. PTMs may be an interesting alternative to the other well-known 3D acquisition techniques, like 3D scanning, which are "expensive" in terms of equipment, acquisition time and post-processing. Moreover, visualization of detailed 3D models is problematic in online environment while PTMs can be a good representation to make data more accessible to the public. Finally, PTMs seem to be a better solution for the visualization of certain objects, like bas-relieves, where the information provided by re-illumination is more important than the one provided by geometry. For all of these reasons our goal is to define a fast, flexible, low-cost system for PTM acquisition of large objects, dealing with all the issues related to the adaptation of the usual "PTM acquisition pipeline" to bigger spaces. Therefore, after an overview of related work in Section 2, in Section 3 we analyze the method of acquisition in all its aspects, from acquisition planning to data processing. In Section 4 we present our solution for browsing of high resolution PTMs, and in Section 5 we propose a preliminary analysis of quality in comparison with 3D scanning, and a study on quality degradation relatively to the number of shots used to calculate the illumination function. Some examples of objects acquired with our system are presented in Section 6. Finally, conclusions and future work are outlined in Section 7.

2. Related work

The appearance of a surface normally varies under different lighting conditions. Traditional CG techniques that involve 3D geometry modeling, specification of reflectance characteristics and lightning conditions, and global illumination rendering methods are used to capture such appearance [DRWP04, Goe04]. But modeling the reflectance function of an object can be very time consuming and computationally intensive. That is why Image-based Relighting (IBRL) can be a fast, alternative method to reproduce photo-realistic lighting effects, like subsurface scattering, interreflection, shadowing and refraction [CC06]. Reflectance imaging is a photographic process that captures views of a surface under varying lighting conditions. For a static object and camera, per-pixel reflectance functions can easily be captured and modeled. Polynomial Texture Maps is a simple but effective technique proposed by Malzbender et al. [MGW01] in 2001. For each pixel, the reflectance function is approximated by a biquadratic polynomial in the following way:

$$\begin{aligned} L(u,v,l_u,l_v) &= a_0(u,v)l_u^2 + a_1(u,v)l_v^2 + \\ &+ a_2(u,v)l_u l_v + a_3(u,v)l_u + \\ &+ a_4(u,v)l_v + a_5(u,v) \end{aligned} \tag{1}$$

where (l_u, l_v) is the direction of the incident light and (u,v) are the pixel coordinates. Hence, each pixel of a PTM image is composed by the RGB values and the six coefficients of the model function. Coefficients are calculated starting from a set of photos taken from a fixed point of view, with different light positions. In order to estimate the coefficients $(a_0, \ldots, a_5)$ the light positions have to be known. Several Cultural Heritage projects used PTM for the inspection of artifacts. The first one was imaging of cuneiform epigraphy [MGWZ00]. HP Labs PTM viewer [MGW01] was used to inspect clay cuneiform tablets under different (optimal) light conditions. Reflection transformation tools were used also in Paleontology, to provide noticeable improvement in imaging of low color contrast, high relief fossils [HBMG02]. The application of PTM method on ancient stone tools revealed fine details of concoidal knapping fractures, use scarring and stone grain [Mud04]. A joint work done by National Gallery and Tate Gallery of London showed that PTM under specular enhancement provided additional information about the surface textures of oil paintings [PSM05]. Cuneiform tablets were analyzed using both 2D (PTM) and 3D (structured light scanner) information. The PTMs were texture mapped on the model, and a special 3D viewer was created [Mud04]. Recently, the application of PTMs and scanning techniques on a large numismatic collection permitted the creation of a "virtual exhibition" [MVSL05]. Moreover, the use of specular enhancement and diffuse gain produced an improvement in data discernment. These brief overview covers the most important applications of PTM in Cultural Heritage; but Polynomial Texture Map can have even other applications, such as investigative work [Mor03] or actor performances [WGT*05].

3. Method

The usual PTM acquisition pipeline had to be re-designed, since the acquisition of object of medium to large size presents several specific issues. In fact, typically PTM are acquired by positioning the object of interest inside a light dome of fixed size (see Section 3.2). This permits to automatically change the light direction during photos acquisition, but limits the flexibility of the overall system. Since the size of our objects is too big to create a fixed dome, we decided to deal with a "virtual" light dome as explained in the next sections. In particular, we divided the whole process in three steps. First of all we considered the physical acquisition. The size of the objects, and the fact that in most cases they cannot be moved from their place led us to the necessity to think about a new kind of non-fixed acquisition system. The proposed system is very simple, but to fasten the acquisition, the planning had to be considered a critical aspect. Acquisition of the photo set was very useful, and some work was done also for the data processing. We present our experience about each step in the next 3 subsections.

3.1. Acquisition planning

Selecting the correct lighting point is an important step in the PTM acquisition of large objects; given the size and position (in the majority of cases, on a wall) of an object, in general we do not have the possibility to use a physical dome to illuminate the object. Instead, we will have to manually place the light in different positions, forming a "virtual" illumination dome. The size of this illumination dome and its light distribution will depend on the size of the target object and on the number of light directions we want to use to sample the reflectance function of the object. To simplify the light placements we developed a software tool, called *PTM Planner*. With this tool it is possible to define the properties of the lighting dome, to visually check its correctness and to automatically generate the coordinates for the light placements. The tool usage is quite simple; the scene setup is generated as the user inputs the size of the object to be acquired, its height from the ground and the distance of the camera. Objects in the scene are scaled according to user specifications; camera is pointed towards the center of the object. Next step is the definition of the acquisition pattern. The array of light can be generated by choosing the light distance and two angles (vertical and horizontal step). The tool can automatically exclude the light positions that are near to the "wall" (there can be problems in positioning the light source in such position) and that are aligned with the camera axis (light will be shadowed by camera or will occlude the camera).

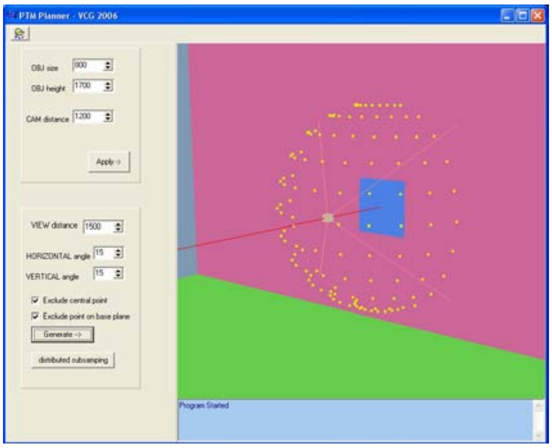


Figure 1: *PTM Planner tool.*

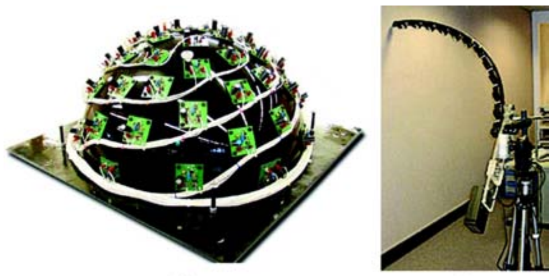
The points are generated using a parallel-meridian grid as showed in Figure 1. This method does not guarantee a uniform distribution over the sphere but, as we will show describing the acquisition procedure (Section 3.2), having a series of light position at the same height will result in a much faster acquisition. For some parameters, it can happen that some points are generated below the floor; those points are automatically excluded. The user can also manually *turn off* (by clicking on the 3D view) the light positions that will probably be impossible to be used due to occlusions. A more advanced possibility could be to automatically detect light position that are impossible to be used due to obstruction caused by the camera tripod (in cases where the camera is between the light and the object). Moreover, a complete description of the room geometry would be required to perform an exhaustive check of the correctness of each light position. However, we believe this level of automatic checking would be too cumbersome, giving a minimal advantage to the user, who is able to evaluate much more efficiently the presence of occluders. Finally, given a complete dome, the program can perform a light pruning following the "distributed" scheme (described in Section 5). This scheme, by generating a more uniform distribution, greatly reduces the number of required light positions while not influencing excessively the PTM quality. When the light setup has been completed, the PTM-Planner tool can save a written description of the scene. The saved data consists in:

- **Scene Data**: a review of the chosen parameters (dimension, angles), total number of photos and instructions for acquisition.
- **Acquisition Plan**: the description of all the points where the light should be positioned during acquisition. The points are described row by row, using cylindrical coordinates: as will be explained in Section 3.2, using an angle template on the floor it is very easy to place the light in the correct position by measuring its distance from the center and its height.
- **PTM Script**: the software that builds the PTM from the photos need to know the light position used for the photos; the PTMPlanner saves the position data in the required format. User will only have to manually change the filename of the images.

Even though PTMPlanner is a quite simple software, it greatly helped us in speeding up the acquisition process both during planning, by giving visual feedback and instant parameters editing, and during acquisition, by providing step-by-step instructions on light placement.

3.2. Acquisition

Several experimental devices has been created to acquire PTMs. Two of them are shown in Figure 2.



(a) (b)

Figure 2: *(a) PTM acquisition dome; (b) PTM acquisition arc (Photos: HP Labs)*

The object in Figure 2(a) is suitable for sampling small objects (nearly 15 cm.). It is a 90 cm diameter black plastic hemisphere, with fifty evenly distributed strobe lights mounted such that they illuminate the hemispheric dome's interior. The digital camera is positioned at the top of the hemisphere and photographs the PTM subject through a view port cut in the dome. In Figure 2(b) a device designed for larger objects is shown. A 90 degree arc 1.50m in diameter is mounted with 12 strobe lights facing towards the center of the arc. One end of the arc is connected to a circular bearing race in the shape of a doughnut. This allows the arc to spin in a 360 degree circle around the bearing race.

These two experimental devices work very well but they are not suitable for our target. Various reasons support this statement:

- The diameter of the "hemisphere" formed by all the light positions depends on the size of the object, since for each photo the light must completely cover the target. For the object shown in Section 5, the minimum diameter for the “virtual” dome was 3 m.
- The lights which compose the arc in the machine in Figure 2(b) can’t completely illuminate a large sized object. We need heavier and more powerful types of light.
- In most cases, the target object cannot be moved from its place, so we have to deal with the fact that it’s not always possible to exploit all the light positions, due for example to the height from the ground.

Figure 3: *Our acquisition setup.*

Following these remarks, we had to think to a simple and cheap acquisition equipment. Our solution is shown in Figure 3.

Since it was not feasible to use a big number of lights, we decided to use only one, and to change its position for every photo of the set. The time needed to position the light was minimized by an accurate planning, as shown in the previous Section, and by some references placed on the floor. We fastened the acquisition using a printed scheme of the angle directions (it helped in placing the references on the floor very quickly), and a plumb line attached to the light in order to facilitate the positioning. Our acquisition equipment was composed of an 8MPixel Canon Digital Camera, a 1000W halogen floodlight, a tripod and a boom stand. The fact that we used only one light explains also the parallel-meridian placement of lights: with these configuration we needed to set the height and direction of the light only once for each level of height. The acquisition can be summarized in this way:

- Take the measures of the object, find the center of it and its height from the ground.
- Using these data, generate the “virtual dome” and choose the positions of all the lights.
- Position the digital camera on the tripod. Measure aperture and shutter speed under the illumination of the central light. Keep these values fixed for all the photos, in order to have a constant exposure.
- With the help of the output of PTM planner, put the reference marks related to each light.
- For each level of height, set the height and the direction of the light, then put it on each reference mark related to the level, and take the photo.

Following this approach and the results of the quality assessment (see Section 5), we were able to acquire several PTMs of an object in a short time (see Section 6). Other big advantages of this equipment are that it is quite cheap (nearly 1000 Euros in total) and easily transportable.

3.3. Data processing

In order to calculate a precise illumination function, a critical factor is that the digital camera must not move from one photo to the other. Even a misalignment of a few pixel can produce a bad result, with visible aliasing. In our experimental acquisition set it’s almost impossible not to have small movements of the camera. This led to the necessity of aligning the set of photos before building the PTM. We performed the alignment automatically using a freeware tool for panoramic images. This was the only data processing we had to add to the usual PTM construction pipeline: there was no need of any image to image calibration, since all the photos had the same exposure. The PTMs were created and visualized with the software we developed. Section 4 will provide a detailed description of how it works.

4. Remote Fruition

In this section we give a brief description of another set of tools that we developed to efficiently browse PTMs of very high resolution.

As previously stated, PTMs are a good multimedia representation of artifacts since the interaction with light in general satisfies the user and its size is typically lower that of a 3D models with fine details. Nevertheless, when the resolution of PTMs is considerable high, such as thousands by thousands pixels, the size of the PTM make it necessary to find a specific solution for browsing it. In fact, a non-compressed PTM of 4000×3000 pixels occupies about 70MB of spaces making a download of it quite tedious. For this reason we decided to develop a specific tool to browse a multi-resolution version of huge PTMs. In this way an Internet user can interact with a low-resolution version of the PTM, while medium and high resolution portions of the image are transmitted when needed, using a progressive transmission approach.

To be more specific, two tools has been developed for this purpose; a first tool, called *HPTM Builder* which gets in input an high-resolution PTM and decomposes it properly into sub-ptms; and the effective browser, a Java applet called *HPTM Browser* (Huge PTM Browser) which handles the stored sub-ptms and browses them following a multi-resolution scheme. Both these tools are based on the Java PTM Library developed by Clifford Lyon (`http://ptmviewer.dev.java.net`).

4.1. HPTM Builder

First of all, a high resolution PTM is decomposed in several sub-ptms in a proprietary format and stored on the server. This sub-ptms represent the leafs of a quadtree, which support the progressive transmission of the level of details of the PTM. The format of these patches is slightly different from the standard format defined by HP. The HP JPEG PTM format treats the coefficients of the pixel as separated images and compress each plane coefficient using standard JPEG/JFIF compression. Before compression the range of values of each plane coefficient are reduced by a scale and bias factor. Our format differs from this one in the compression algorithms used: in place of JPEG/JFIF compression we use lossless JPEG2000 compression. We have not studied in deep the consequence of using lossy JPEG2000 compression. This aspect is left as an interesting work for future investigations.
During the decomposition each sub-ptms is properly indexed to be retrieved efficiently by the PTM browser in a multi-resolution way. In particular, chosen the number of resolution levels of the quadtree the corresponding indexed sub-ptms are generated. The indexing scheme is done accordingly to the Z-order space filling curve. This curve is particularly suitable for this purposes since it is easy to convert from the d indices of a d-dimensional matrix to the 1D index along the curve. For further information about this topic see [PF01, BKV99] In other words it is simple to navigate through the quadtree without storing it explicitly by using a matrix, filled using the Z-order filling curve, that we call Z matrix. For example, two levels of resolution produce 16 sub-ptms indexed by the following Z matrix:

$$Z_2 = \begin{bmatrix} 0 & 1 & 4 & 5 \\ 2 & 3 & 6 & 7 \\ 8 & 9 & 12 & 13 \\ 10 & 11 & 14 & 15 \end{bmatrix} \qquad (2)$$

where $0,4,8$ and 12 are the node at the coarse level of resolution, $1,2$ and 3 are the child of the node 0, $5,6$ and 7 are the child of node 4, and so on. It is possible to exploit the property of this matrix to efficiently retrieve the patches at a given level of resolution. For example, as it is possible to notice, the indices of all the patches at a level of resolution l differs from $\Delta = 4^l$.

4.2. HPTM Browser

The *HPTM Browser* is a Java applet composed by two panels, a view panel where the user can interact to change the light and explore the image, and a navigation panel that shows the part of the PTM currently under viewing with respect to the whole PTM. During the interaction a thread called *update thread* loads from the server (using the HTTP protocol) the sub-ptms increasing progressively the level of details of the part of the PTM under viewing. The update thread works with a resolution table T that stores the information about the patches loaded. In fact, all the patches have the same resolution, but when a parent node is loaded it is interpolated to cover all the part of the PTM at the target level of resolution. When a level of resolution is complete the thread starts to load the sub-ptms corresponding to the immediate child patches. The process ends when all the leafs are loaded.
It is important to underline that to obtain optimal performance the update thread processes only the set of patches currently under viewing. The set of patches under viewing is calculated whenever the user pan or zoom the PTM. In this way the browser is capable to visualize very quickly a PTM at full resolution after a big zoom operation. Obviously, obtaining the whole PTM at full resolution requires a long time since all the sub-ptms have to be download.

5. Quality assessment

Besides trying to create high resolution PTMs of large artifacts, we wanted also to consider some issues regarding quality assessment. In particular, we performed some quality evaluation both with respect to the "best" representation of the object and the number and position of lights considered for acquisition. In order to perform this quality evaluation, we considered a 70 by 80 cm section of the XIVth Century Tomb of Archbishop Giovanni Masotti as a case study. We performed a very accurate PTM acquisition, using a large number of lights (105 light positions, 11 angles and 11 height levels). We acquired the same object also with a triangulation Scanner (Minolta 910i). We consider the 3D scanned model as a "ground truth". For larger objects 3D scanning represents a very reliable technique, in terms of accuracy [BR02], even if recently it has been proposed to improve its field of normals using a combination with photometric stereo [NRDR05]. Following the pipeline described in Section 3, we created a PTM using all the 105 photos. We also generate an high-precision 3D model (nearly 2.4 millions of faces, $\frac{1}{3}$ of millimeter of sampling resolution) from a set of 68 range maps. Our first comparison was between these two representations, in order to estimate the quality of the normals calculated from the PTM data. We aligned the 3D scan model to the PTM [FDG*05] and we calculated the normals of both the model and the PTM. In Figure 4 a comparison of the normal maps is shown. The variation of the normals in the PTM is smoother than in the corresponding 3D scan, but their values are coherent. This test demonstrates that, even though PTM provides an approximation of the objects' geometry, the obtained data are reliable. It also demonstrates that our setup doesn't introduce errors in the representation. The other analysis was related to the degradation of PTM quality respect to the number and position of lights. For this purpose, we created four PTMs starting from subsets of

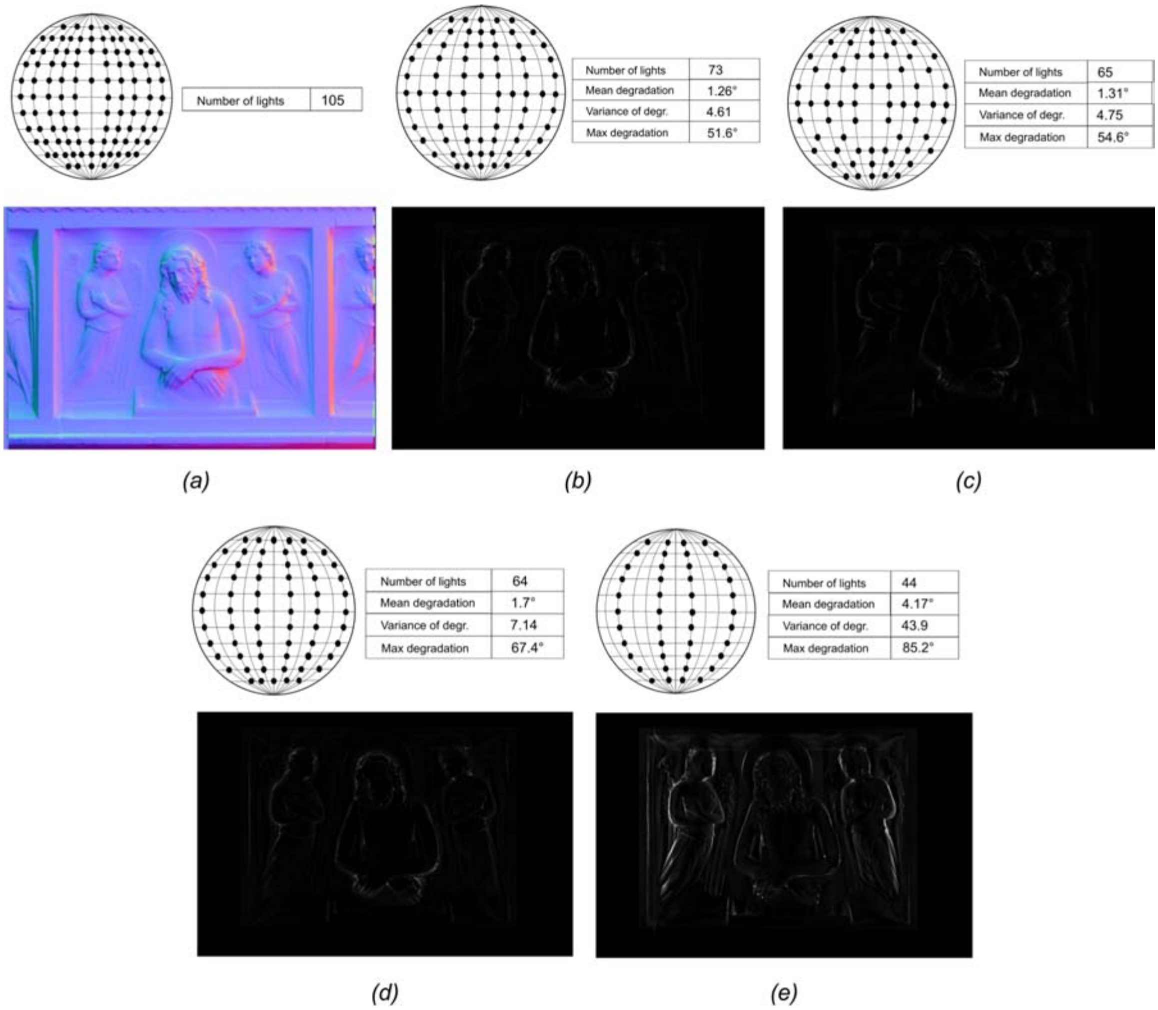


Figure 5: *Quality degradation: (a) Best quality PTM (normal map) (b-e) Maps of the differences in dihedral angle of normals. The sphere shows the lights placement.*

the original lights. Then we made a comparison between the normal maps of the "best" PTM (the one with 105 lights) and the "subsampled" ones. The comparison was made calculating the difference in dihedral angle between the normals of each pixel. In Figure 5 we show the analysis of the difference between the best PTM and four possible subsets. In terms of number of lights, we can observe that we can considerably reduce the number of lights without having an excessive degradation of quality. For example, we can reduce the number of photos up to 65 (see Figure 5(c) and 5(d)) and we will have a PTM where mean value and variance (nearly 1.5 and 6 respectively) of the overall degradation are still satisfying. As regards the different placement of lights, we can observe the case of Figure 5(c) and 5(d). Even though we have almost the same number of lights, a more "distributed" position of the lights brings to lower mean degradation and peak error. Considering these facts, we can conclude that a set of 60-70 properly distributed photos can produce a high-quality PTM. The results of this quality assessment were used to reduce the acquisition time for the objects shown in the next section.

6. Results

Several objects have been acquired with the developed system. In this Section we present 3 examples of the PTMs we produced; all this material is available for real-time exploration at `http://get-me.to/ptm` . The first example is one face of a small (30 × 30 × 20 cm.) Medieval Capital from the Museum of S.Matteo in Pisa. With the help of a professional photographer, we created a set of 36 high resolution (5440 × 4080) photos. In this case, we did not

Figure 4: *Comparison between the normal maps of the 3D scanning and the PTM: full model and particular.*

use the equipment described in Section 3.2, but a 20 MPixel Monorail View Camera and a professional flash light. We

Figure 6: *The Museum of San Matteo Capital.*

produced a very detailed horizontal PTM of the Capital: a snapshot is shown in Figure 6. The acquisition time for this object was nearly 1 hour.

The second example was already considered in Section 5: it was a part (70 × 80 cm.) of the XIVth Century Tomb of the Archbishop Giovanni Scherlatti, by Nino Pisano (Museum of the Opera Primaziale in Pisa). We performed a very detailed acquisition (105 light positions, image resolution 3496 × 2280), which lasted about 3.5 hours. In this case we used the acquisition system described in Section 3.2. We were able to produce a very detailed PTM: in Fig 7 we show a snapshot. The third test object was a II Century A.D. Roman Sarcophagus, representing the Phedra and Hyppolitus Myth. This artifact is in the Camposanto Monumentale of Pisa. We chose this particular example for two main reasons: first of all we wanted to make use of the results of the quality assessment in order to perform a detailed acquisition with a lower

Figure 7: *Placed Christ from the Museum of Opera.*

number of lights (and consequently a shorter time). The sec-

Figure 8: *Roman Sarcophagus from Camposanto Monumentale.*

ond reason was that the Sarcophagus is situated outdoors, so we wanted to experiment if the proposed acquisition system could produce a good PTM also when the ambient light is considerable high with respect to the light equipment we used.
After 2.5 hours of work we were able to produce a complete PTM of a 90 × 60 cm portion (66 photos, resolution 3496 × 2280, see Fig. 8), and two horizontal PTMs (10 photos each) of the two halves of the Sarcophagus As can also be seen in the web site, the results was more than satisfying, considering the non ideal condition of lighting.

7. Conclusions and Future Work

In this article we extended the use of Polynomial Texture Mapping to large objects, and we demonstrated that it is possible to produce very detailed representations with a low-cost and simple acquisition system. In order to achieve this result, we had to rethink the PTM acquisition pipeline. Due to the impossibility to utilize the usual fixed acquisition

dome, acquisition planning became very important. We defined also a quite simple acquisition system, and we analyzed the data processing step. After an accurate planning the acquisition can be very fast, and the equipment needed is cheap and easy-to-use. In order to preserve the high detail and the interactivity of exploration, we developed new software to progressively browse PTMs. Moreover, we studied some issues about quality assessment of PTMs. This studies gave us useful suggestions to perform the acquisition more quickly, without losing quality in the final result. All the examples gave satisfying results, and showed us that PTM can definitely be an alternative method for documenting and communicating Cultural Heritage information also for large size objects. In particular, the variation of illumination can be the best way to inspect objects such as bas-relieves or paintings.
Some future work could be useful to improve the technique. We can exploit an automatic system to estimate the light direction starting from the photo set in order to make the acquisition process more fast and accurate. Another useful feature could be the removal of the ambient lighting contribution, which lowers the quality of representation especially for outdoors objects.

Acknowledgements

We would like to thank the museums of San Matteo and Opera Primaziale in Pisa for their support. We would also thank Mark Mudge for his suggestions and enthusiasm about PTM technology. We also thank Clifford Lyon for share with us and with the Java community his Java PTM library. This work was partially funded by Epoch: European Research Network on Excellence in Processing Open Cultural Heritage (EPOCH) and Cultural Heritage Department of Italian National Research Council.

References

[BKV99] BALMELLI L., KOVACEVIC J., VETTERLI M.: Quadtrees for embedded surface visualization: Constraints and efficient data structures. In *Proceedings of International Conference on Image Processing 1999* (October 1999), vol. 2, pp. 487–491.

[BR02] BERNARDINI F., RUSHMEIER H. E.: The 3d model acquisition pipeline. *Comput. Graph. Forum 21*, 2 (2002), 149–172.

[CC06] CHOUDHURY B., CHANDRAN S.: A survey of image-based relighting techniques. In *Proceedings of the 2nd International Conference on Computer Graphics Theory and Applications* (March 2006).

[DRWP04] DEBEVEC P., REINHARD E., WARD G., PATTANAIK S. (Eds.): *High Dynamic Range Imaging, (ACM SIGGRAPH Course Notes No.13)*. ACM Press, 2004.

[FDG*05] FRANKEN T., DELLEPIANE M., GANOVELLI F., CIGNONI P., MONTANI C., SCOPIGNO R.: Minimizing user intervention in registering 2d images to 3d models. *The Visual Computer 21*, 8-10 (sep 2005), 619–628. Special Issues for Pacific Graphics 2005.

[Goe04] GOESELE M.: *New acquisition techniques for real objects and light sources in computer graphics*. PhD thesis, Universität Saarbrücken, 2004.

[HBMG02] HAMMER O., BENGTSON S., MALZBENDER T., GELB D.: Imaging fossils using reflectance transformation and interactive manipulation of virtual light sources. In *Paleontologia Electronica* (2002), no. 1.

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

[MGWZ00] MALZBENDER T., GELB D., WOLTERS H., ZUCKERMAN B.: *Enhancement of Shape Perception by Surface Reflectance Transformation*. Tech. Rep. HPL-2000-38R1, Hewlett-Packard Laboratories, Palo Alto, California, 2000.

[Mor03] MORTON M.: Foresinc control examiner for the san francisco police criminalis laboratory, presentation to the criminalistics section of the american academy of foresinc sciences, 2003.

[Mud04] MUDGE M.: Siggraph 2004 conference presentations, web graphics/special sessions/ panels, cultural heritage and computer graphics panel, 2004.

[MVSL05] MUDGE M., VOUTAZ J.-P., SCHROER C., LUM M.: Reflection transformation imaging and virtual representations of coins from the hospice of the grand st. bernard. In *Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2005)* (November 2005), Mudge M., Ryan N., Scopigno R., (Eds.), Eurographics Association, pp. 29–39.

[NRDR05] NEHAB D., RUSINKIEWICZ S., DAVIS J., RAMAMOORTHI R.: Efficiently combining positions and normals for precise 3D geometry. *ACM Transactions on Graphics (Proc. SIGGRAPH) 24*, 3 (Aug. 2005).

[PF01] PASCUCCI V., FRANK R. J.: Global static indexing for real-time exploration of very large regular grids. In *on-line proceeding* (November 2001).

[PSM05] PADFIELD J., SAUNDERS D., MALZBENDER T.: Polynomial texture mapping: a new tool for examining the surface of paintings. ICOM Committee for Conservation, 2005.

[WGT*05] WENGER A., GARDNER A., TCHOU C., UNGER J., HAWKINS T., DEBEVEC P.: Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Trans. Graph. 24*, 3 (2005), 756–764.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Current Practice in Digital Imaging in UK Archaeology

Alice Chuter and Kate Devlin

Department of Archaeology and Anthropology, University of Bristol, UK

Abstract

The field of archaeology relies heavily on photography as a way of recording information about sites and artefacts. It is therefore essential that we can have confidence in the photographic record, as any changes over time can result in information being lost forever. In the past five years digital imaging has become a potential alternative to traditional film photography. However, this has great implications, as both methods have very different advantages and disadvantages. Most notably, if the rise in digital photography in the heritage sector mirrors that of current public trends, there is a danger that digital preservation issues could be neglected. We undertook a survey of photographic practice among archaeologists in the UK in order to gain an insight into the prominence of digital photography for recording our past. This paper presents and analyses our results.

Categories and Subject Descriptors (according to ACM CCS): J. Computer Applications [J.2 Physical Sciences and Engineering]: Archaeology

1. Introduction

Photography is an essential part of recording information about archaeological sites and artefacts. For around 150 years this has been in the form of chemical (silver halide) film photography, but in the last five years digital imaging has become an alternative to traditional film-based methods. However, differences in these media have great implications. There is uncertainty as to whether those working in cultural heritage are using digital technology only after careful considerations of its advantages and disadvantages, or if they are simply following current public trends.

Photographs used in the recording of excavations and artefacts are by necessity of a high quality, both in the prints themselves and in the skills of the photographer. If a change in technique was to influence this adversely this would be a huge loss to the discipline. Digital photography offers advances such as convenience, immediacy, ease of dissemination, and long-term savings in equipment costs. However, the use of film photography still generally offers better quality and longer lasting images. Current debate in the field of digital archiving shows that often people are unaware of issues surrounding digital preservation, or are unable or unwilling to address the changes that digital imaging brings about.

We undertook a study of photographic practices in UK archaeology. The results gave an detailed insight into current practices and views on the rise of digital imaging in the cultural heritage sector, specifically in archaeology. This paper presents the findings and analyses the current attitudes towards recording the archaeological record heritage in a digital format.

2. Background

Although digital imaging – the capturing of images via an electronic rather than chemical means – has been around since 1981, it is only relatively recently that there has been a rise in the mainstream use of this technology [Gre05]. Compared to the 150 year old history of traditional film photography, the digital technology involved is new and the full implications have not yet been fully assessed. In UK archaeology the changes have begun – the decision making process is no longer ‘large format or 35mm?’ but ‘digital or film?’ [Hou06].

The photograph is ultimately a descriptive medium, one which is said to speak a thousand words. It can relay the complex detail of its subject, show the relation with other things present and depict a precise location. One of the key roles of photography in archaeology is to capture a moment of time: a series of photographs are taken over the duration of an excavation, for example, and thus capture each stage, producing a kind of timeline. Photographs portray specific factual information, and this is important in archaeological recording where an attempt is usually made to strip images and plans of as much subjectivity as possible. This is not to say that photographs are completely objective. The photograph itself is the product of the photographer and there are always some choices to be made that will affect the picture, such as where the photographer stands or how high or low they hold the camera [Cla97]. With the rise of digital photography, altering photographs is becoming much easier, it is possible that the relative objectivity of archaeological photography could be seen to decline.

Photography was invented in the mid 1800s, beginning as the basic concept of capturing a negative image on silver chloride-impregnated paper which was then contact printed to give a positive image. The English patent-holder of this process, W.H. Fox Talbot, was also one of the first antiquarians to record archaeological finds with photography [Dor94]. Photography was an important contributor to the shaping of archaeology as a more scientific-based discipline with an increased analytical approach. From the end of the nineteenth century photography became a standard technique for recording excavations as an accompaniment to illustration, and some of the early photographs are still in a good state of preservation some 100 years later [Dor94]. Today, the use of photography in archaeology ranges from aerial photography, infra-red photography for environmental analysis, photogrammetry, data recording and public relations, to name a few. Our investigation concentrates on photographic practice by UK archaeologists engaged in fieldwork.

3. Method

We sought to investigate opinions in UK archaeology regarding digital imaging. We wanted to gain an insight into the uptake of digital photography in archaeology, the attitudes towards it, whether the type of photography was influenced by purpose, and whether people were aware of digital preservation issues. We designed a questionnaire with an easy-to-read mixture of multiple choice and open-end questions. It consisted of twenty questions split into four sections: General, Personal Opinions, Film and Digital. The questionnaire is shown in Table 1. An email survey was deemed most appropriate, with the questionnaire placed in the body of the email. Ninety-two questionnaires were sent out to addresses in the Council for British Archaeology's (CBA) email directory – twenty-six to UK universities and sixty-six to archaeological contractors and units. Eighty-four of these were successfully delivered.

4. Questionnaire Results

Of the eighty-four email questionnaires sent out, there were forty detailed replies, representing a 49% response rate. The results that follow are broken down by questionnaire section.

4.1. Section 1: General

The survey showed that 95% (37/39) of archaeologists use both film and digital in their work despite current consumer trends where the majority of cameras sold to the public today are digital, illustrated in the way that corporations such as Kodak, Nikon and Konika Minolta are switching entirely from film to digital imaging. The main reasons given for using film were archive stability, quality of results, and tradition, whilst the main reason for using digital was ease (for example manipulation, quick emailing capabilities, and ease in inserting photographs into lectures and reports). Other reasons for using digital photography included the ability to verify the image immediately on site, flexibility and versatility, and low cost. However, 32.5% (13/40) of respondents stated that the reason they used a particular system of photography was due to rules imposed by curators, museums and county archaeologists. Only one respondent stated that they used digital photography exclusively, stating 'digital suits the requirements of the work best and I see no need at all to do film photography any longer'.

The responses showed that 60% (24/40) of respondents use different systems of photography for different projects, whereas 22.5% (9/40) do not, 10% (4/40) stated it would depend on circumstances (such as how quickly the image was needed, the resolutions required, and the purpose of the picture), and 7.5% (3/40) answered that it was the decision of the curator. As illustrated in Figure 1, the general consensus was that if a photograph can be taken again (for example artefact recording or surveying), digital is often favoured, whereas if the image cannot be taken again (for example, during excavations), film, or a combination of both film and digital, is preferred. This appears to reflect the opinion that digital photography is less reliable.

With respect to purpose, the answers to the questionnaire showed that the majority of photographs are used for publication (27%) and teaching (20%), with archives (17%) and site records (14%) following (Figure 2). Publication work and archives often use traditional photography, whereas teaching and site record photographs tend to favour digital, showing that digital and traditional photography seem to have an average split.

Regarding the preservation of photographs, the results show that 85% (33/39) of respondents store their photographs in both an analogue and digital fashion. Analogue storage includes prints, slides and negatives, whereas digital storage includes hard drives, servers and CDs. The remaining 6/39 respondents were divided equally between analogue and digital storage.

Archaeological Photographic Survey

General:
1) What system of photography do you use in archaeological projects?
Digital / Film / Both
2) Why do you use that system (for example cost issues, easier use of equipment and processing, etc)?
3) Why do you NOT use an alternate system?
4) Do you have a different system for different projects? For example, sites/artefacts photography?
5) What are your photographs used for (for example, publication, records, slides, etc)?
6) How do you store your photographs?
Prints / Digital / Both
7) Do you have a professional photographer on site/ taking artefact photos?

Personal Opinions:
1) Which system do you think prints better photographs?
Film / Digital / Equal
2) Do you think digital manipulation is an advantage or disadvantage?
Advantage / Disadvantage / Not really thought about it
3) What is the main reason for the above answer?

Film:
What type of film do you use?
Colour / Black and White / Slide / A combination (please specify)
2) What equipment is necessary apart from the camera itself?
3) How easy is it to look after the camera and equipment on site (for example, carrying it around, preventing it getting dirty)?
4) How easy is to adjust the settings on your camera?
Easy / Can be difficult occasionally / Difficult / Other (please specify)

Digital:
1) Are batteries running out a problem on site?
Yes / No
2) What equipment is necessary apart from the camera itself?
3) How easy is it to look after the camera and equipment on site (for example, carrying it around, preventing it getting dirty)?
Easy / Can be difficult occasionally / Difficult / Other (please specify)
4) How easy is it to adjust the settings on your camera?
Easy / Can be difficult occasionally / Difficult / Other (please specify)
5) Do you transmit photos online straight from a site?
Yes / No
6) Do you find the resolution of digital prints a problem?
Yes / No

END

Thank you for taking the time to fill out the questionnaire.

Table 1: *Questionnaire used in this study.*

The last question of the ‘General’ section revealed that 75% (30/40) of archaeological excavators do not have a professional photographer on site, 15% (6/40) sometimes do, and 10% (4/40) do so permanently.

4.2. Section 2: Personal Opinions

When asked which system prints better photographs: film, digital or equal, 40% (16/40) of recipients answered ‘film’, 10% (4/40) answered ‘digital’, 30% (12/40) thought they were equal and 20% (6/40) said that that the quality depends on a certain factor. The factors they suggested included the quality of the digital camera (the higher the pixel value the better), quality of the printer, paper and ink for digital printing, the skill and experience of the photographer, the type and quality of film used in traditional cameras, and the processing lab used. However, although the quality of digital prints is approaching the quality of film prints, the latter currently still has the edge.

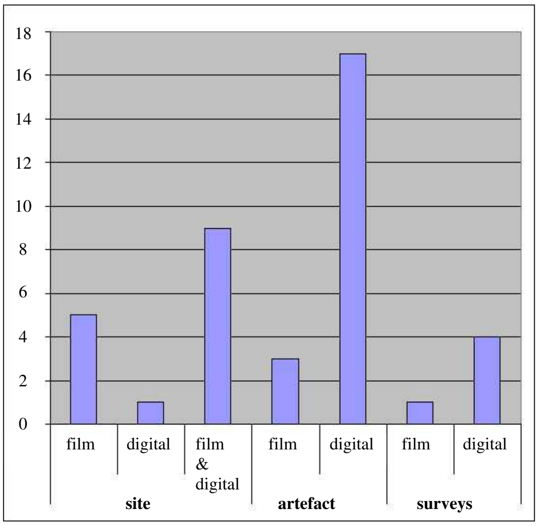


Figure 1: *Breakdown of positive responses to the question'Do you use a different photographic system for different projects?'*

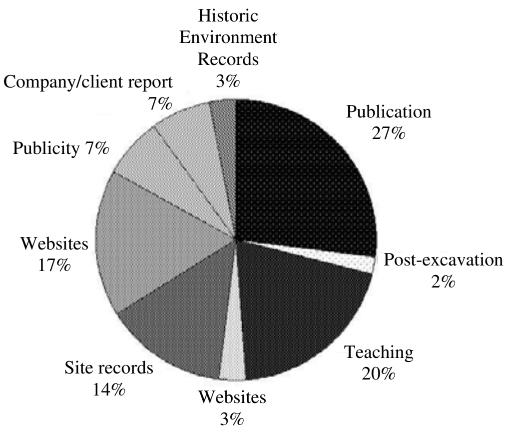


Figure 2: *Pie chart showing uses of photographs.*

A large proportion of respondents, 87.5% (35/40) stated that they thought digital manipulation of photographs was an advantage, 7.5% (3/40) thought it was both an advantage and a disadvantage, and one stated that it was a disadvantage, with one other not having considered this as an issue. However, comments such as 'not always appropriate', 'open to abuse' and 'a code of practice may be eventually required' were stated on many replies, showing that many archaeologists are still wary and possibly dubious about the use of digital manipulation of archaeological photographs. The main reasons for digital manipulation being an advantage were: the ability to 'tidy up' images (for example cropping or zooming to define specific features), the ability to adjust the light levels and colour balance of the photograph, the ability to easily add text, arrows and annotations, and the ability to improve the quality or enhance poor quality photographs. The main reason for digital manipulation being a disadvantage was that it could be misleading. Although only 1/40 recipients stated digital manipulation as a disadvantage, the concern about misrepresentation was in fact mentioned 5/40 times. 7.5% (3/40) of recipients also stated that manipulation of film is also possible.

4.3. Section 3: Film

The first question of the 'Film' section asked about the type of film used. Most recipients used a combination of films, the most popular being slide (42%), then black and white print (36%), and lastly colour print (22%).

Question 2 was the first of the three questions repeated in the 'Digital' section (see 4.4) asking about necessary photographic equipment (excluding the camera itself) used on site. The items were, in order of popularity: tripod, flash, lenses, filters, films, light meter, light box, cleaning equipment/wipes, and batteries.

When polling the respondents on the care of a film camera and related equipment on site, 16/37 found it difficult, 15/37 found it difficult occasionally, and 6/37 found care to be easy. The reasons suggested for difficulties in looking after the cameras included dropping them, cameras accidentally becoming buried or crushed (including under heavy machinery), and getting them covered in rain and mud. However, one respondent stated that older cameras can take more punishment, and 4 out of the 6 of the respondents who gave the answer 'easy' stated that this was only when a camera case and common sense were employed. As one respondent stated, it is technically not difficult to look after a camera on site, but in reality people often neglect to clean them or look after them properly.

Regarding the use of the cameras, 84% (32/38) respondents stated that the settings were easy to adjust on their film cameras, and only 8% (3/38) stated that it was difficult occasionally. No one chose the 'difficult' response to this question, however 8% (3/38) did state that it has more to do with experience, suggesting that it was 'easy if you know how, hard if you don't'.

4.4. Section 4: Digital

One major difference between film and digital cameras is the requirement of digital cameras to run on batteries. Therefore the questionnaire asked whether batteries running out on site was a problem. 67% (26/39) stated that this was not a problem, 31% (12/39) stated that it was a problem, whereas 2% (1/39) stated that it was a problem occasionally. Various solutions to this problem were using re-chargeable batteries, carrying spares, and using in-car chargers. One respondent stated that people were often better at remembering batteries when using digital cameras as opposed to film.

The necessary equipment required for digital cameras on site were in order of popularity: tripod, batteries, laptop, charger, flash, software, memory cards, lenses, cleaning equipment, and a downloading cable. 20.5% (8/39) of respondents stated that the same items were required for digital as for film, and 10% (4/39) stated that no extra equipment was required for digital photography. The list illustrates that the extra equipment necessary for digital photography mainly centres around downloading images, rather than items directly to aid taking photographs.

When asked about camera and equipment care, comments made in response to this question included the need for extra care when looking after digital cameras as they are more expensive to replace. There were suggestions that digital cameras are easier to look after as the body does not need to be opened to access film, but conversely that digital cameras are more difficult to look after as they are less robust than film cameras.

76% (26/34) of respondents stated that the settings were easy to adjust on their digital cameras, 18% (6/34) stated that it was difficult occasionally, and 6% (2/34) stated that it was difficult. Reasons for it being difficult were stated as being due to the newer technology, being difficult to see the LCD screen in bright light, and complex menus on the cameras. One respondent also stated that digital cameras are more likely to have 'fiddly' controls than film cameras.

82% (32/39) respondents stated that they do not transmit digitally downloaded photographs straight from the site, 8% (3/39) stated that they do, and 10% (4/39) stated that they occasionally transmit photographs straight from the site. 5% (2/39) stated that they would consider transmitting photos from the site, but the main reason for not doing so was the lack of suitable facilities on site (such as internet capabilities). One respondent, however, stated that they have sent photographs via camera-enabled mobile phones if unexpected elements have arisen on site.

Regarding the resolution of digital images, 53% (20/38) of respondents stated that they do not find the resolution of digital prints a problem, whereas only half this amount, 26% (10/38) stated that they do. 21% (8/38) of respondents stated that the resolution of digital prints can sometimes be a problem, and that often digital prints are not quite good enough for some publications.

5. Discussion

Four main issues arose from our findings: first, convenience; second, image manipulation; third, skill; and fourth, preservation.

5.1. Main issues

Digital photography in archaeology seems to be mainly used for two reasons; speed and ease. The immediate digital format of the photographs means there is no waiting time for photographs to be developed, therefore making it fast and simple to insert pictures into reports, on to websites, and into presentations – a major advantage for desk based assessments and evaluations, where quick reports are often needed. Alongside this is the ease of digital pictures over film when distributing multiple copies to clients. One respondent gave the example of being able to tie the digital photographs of a specific site immediately into an object database, making the task of recording several thousand objects in an area much easier, faster, and more comprehensively than if film photographs had been used. However, it is also apparent from the results that digital photographs appear to be used for less academic purposes than film photographs. The digital images taken were working shots, website shots, photographs for monitoring sites, for publicity and for general information. Although digital photographs are used in written reports, the general opinion seems to be that film photographs ensure quality, especially in printed publications, whereas digital photographs ensure ease and speed.

A second issue that was raised more than once was the manipulation of photographs. Although many respondents agreed that this was an advantage, there were definite concerns about the subject. Unlike some other areas of photography where artistic license is acceptable or desirable, archaeological photography aims to be an accurate representation and record of what is present at the time, and the integrity of the record is paramount. There are, of course, reasons unrelated to image manipulation as to why photographs can be more subjective than objective; camera angle, lighting and framing, which are all controlled by the photographer and can influence the way that the subject is seen [Sav97]. However, it is the manipulation of the image itself which can cause it to be false and unrealistic, something of great concern in archaeology. Although digital manipulation can be very advantageous when 'tidying up' images (for example cropping, zooming and adjusting colour balance and brightness to a certain degree), the potential to take it too far is evident. As one respondent stated, digital manipulation of images 'opens up a can of worms when considering that your image is supposed to be an accurate representation of the truth – airbrushing out a stray trowel is only the start of a slippery slope'.

One might argue that manipulation of photographs has always been possible, even for film photographs, and this is indeed true. However the difference, and concerns, with digital manipulation is that it can be done quickly and easily; most people can easily learn how to use software suitable for altering images. With electronic computer based methods of manipulation, processes such as combining and blending images from different sources, repositioning features and objects, or even removing them altogether, can be done easily and with such skill that, unlike film manipulation, it is very often impossible to detect. Methods suggested in the results of the survey to avoid this problem are simple yet effective: document altered images, keep the original file format where appropriate, be clear about why the image is being altered, and incorporate rules such as manipulation being acceptable for presentations, but unacceptable for archives. Codes of practice do exist and would be a great advantage if made prominent this discipline.

The third main issue is that of photographic skills and quality. The ability to easily manipulate images can, unfortunately, compensate for bad photography, or as one respondent stated, 'get you out of jail, so to speak'. Many respondents mentioned that it can be difficult to rely on staff to take good photographs, and that they often have limited knowledge. Although this may not necessarily be true for all archaeological units, the results of the questionnaire showed that it is definitely a problem for a significant number of establishments, with replies such as 'training in photography is non-existent', 'increasingly staff have no knowledge of SLR cameras', 'insistence on using manual settings when they [staff] do not know the basics' and 'site photography when completed by site staff, rather than a professional photographer, can be of low quality' appeared frequently enough to make the problem apparent. The results of the questionnaire showed that only 10% (4/40) of archaeological units surveyed employ a full time professional photographer. The main reason for not doing so was cost. When the market today offers such easy access to photography through high-quality compact digital cameras it is possible that there is a reliance on the technology to automate the necessary photographic procedures, with less of a requirement to learn the appropriate manual skills. Simply taking into consideration aspects such as leading lines, balance, and centre of interest, can improve a photograph; making the subject easier to discern but without changing the image itself [Sim69]. Skills such as these can be, to a certain extent, disregarded with digital photography, as manipulation of the image at a later date could rectify them. As Parkhouse [Par05] states in a staff training manual, 'when you look through the viewfinder, look at the edges of the view as well as the subject of the photograph – you may be surprised how many extraneous bits of litter/piles of loose spoil/pieces of clothing/tools there are'. Simple skills like these which have always been taken into consideration in film photography may be considered less important in digital photography, with the photographer safe in the knowledge that they can erase any unsightly and irrelevant aspects of the picture later on. Some respondents even mentioned the reluctance of staff to use a simple but important piece of kit: the tripod.

The final issue is that of preservation. Not all respondents were aware of the need for digital preservation, and although one respondent stated that the 'problem of digital archiving will soon be resolved when the national archives start providing server storage and data migration for site archives', there are other issues involved. In UK archaeology digital storage in physical form generally means either on CD or DVD, internal hard drive or external hard drive. Small units and contractors frequently do not have the money, resources or expertise to regularly back up data onto magnetic tape. There are two main issues concerning digital preservation: media holding digital data being subject to physical decay, and the technology used to retrieve that data becoming obsolete, although the problem of obsolescence is a greater problem than that of media fragility [Tec02]. CD-R and DVD-R are the most common physical digital preservation method employed by those surveyed, although possibly the worst choice, as over the last few years there have been concerns with lost data and unreadable disks [APE*06]. The loss of data stored on this type of media is most likely due to poor storage or writing (burning the data to the disc), and so in many ways does not differ from analogue storage in that it requires sufficient care and attention.

A survey by the Archaeology Data Service (ADS) in 1999 that most archaeological digital information was held on floppy disk, and little use was being made of CDs, as well as a sizable minority of organisations presuming that they would not have the capacity to work with the Internet [CRRW99]. Only six years on we take such things for granted, showing how quickly technology changes. There are of course methods which can be employed to avoid such situations, making digital preservation an active task: monitoring of data, as well as the software and hardware in which they are contained, for possible obsolescence and degradation and a procedure for the appropriate action to be carried out to overcome these problems. Alongside this a risk assessment of the image content is necessary, which means an understanding of the content, how it is held and the risks of obsolescence associated with the format and media [APE*06, PV05]. However, the cost of such active intervention across the life-cycle of a digital archive is unclear, as any existing operational digital preservation systems are too recent to provide stable information. With many UK archaeological contractors operating as small businesses, it seems unlikely that they would be able to meet the potentially prohibitive costs. Fortunately, digital repositories are available through organisations such as the ADS, and our survey indicates that print archives are indeed in place, and continue to be recommended by professional bodies such as the Institute of Field Archaeologists.

5.2. Other issues

The cost of digital cameras ranges vastly, with many being cheaper than film cameras, but high resolution digital being much more expensive. Assuming that the resolution of 35mm film is approximately the same as a digital resolution of 12-20 megapixels [Woo04], when comparing the price of an SLR film camera to a comparable resolution digital camera, the digital camera is currently over double the price of the film camera. The results of the survey indicated that the most of the respondents were using digital cameras of around 6 megapixels resolution, and several respondents complained of the cost of digital imaging equipment. However, the results of the questionnaire showed that just over half of the respondents do not find the resolution of digital prints a problem.

Wooliscroft [Woo04] suggests that an advantage of using digital cameras on site is the ability to download photographs on site and subsequently transmit them to conservation specialists or heritage bodies. Although this sounds advantageous, the results of the questionnaire show that in reality few archaeology units do this, with only 18% of respondents transmitting or sometimes transmitting pictures from site. Not unexpectedly, the reason for not doing so is mainly due to lack of suitable resources on sites such as internet connections or even electricity.

In general, the results of the survey indicated that respondents found settings harder to adjust on digital cameras, with 24% of respondents choosing 'difficult' or 'sometimes difficult', as opposed to only 8% for film cameras. Comments from respondents suggests that this is due to complicated interfaces. Keeping a copy of the camera manual to hand was mentioned by various respondents only for digital cameras, not for film cameras, and one responded stated that 'the basics of camera operating are more easily demonstrated on traditional cameras'.

5.3. Conclusions

Overall, our survey suggests that although digital imaging has become the mainstay of the consumer market, UK archaeologists have approached it with consideration and are not just 'following fashion' as suggested by Wooliscroft [Woo04]. Film photography is still widely used and is for the most part regarded as more stable in terms of quality, reliability and longevity. However, digital imaging has found a niche amongst archaeologists with definite advantages such as convenience and ease of use.

Somewhat worrying, however, is that not all those questioned were aware of the need for digital preservation, and there was little to no sign of digital preservation policies actually being implemented. Establishing a secure digital preservation strategy is paramount if archaeological images are to be successfully accessed in the future. This is particularly important for excavation photographs where the archaeological process itself is destructive and the image is an essential record of the past.

Finally, the responses to our survey raise awareness to the fact that digital imaging may be accelerating the decline of photographic skills and knowledge, as poor photography can be rectified with digital manipulation a later stage. We strongly advocate the training of staff and the use of existing standards and guides to good practice regarding digital photography in order to ensure that our cultural heritage can benefit from advances in technology.

References

[APE*06] ANDERSON S., PRINGLE M., EADIE M., AUSTINE T., WILSON A., POLFREMAN M.: *Digital Images Archiving Study*. Joint Information Systems Committee, 2006. URL=http://www.jisc.ac.uk/index.cfm?name=imagesarchiving, accessed 20/04/06.

[Cla97] CLARKE G.: *The Photograph*. Oxford University Press, 1997.

[CRRW99] CONDRON F., RICHARDS J., ROBINSON D., WISE A.: *Strategies for Digital Data: findings and recommendations from digital data in archaeology: A survey of user needs*. Archaeological Data Service, 1999.

[Dor94] DORRELL P.: *Photography in Archaeology and Conservation*. Cambridge University Press, 1994.

[Gre05] GREGORY P.: Digital photography. *Optics and Laser Technology 38* (2005), 306–314.

[Hou06] HOULDER E.: Think before you click. *British Archaeology 87* (2006), 30–33.

[Par05] PARKHOUSE J.: Archaeological photography. Unpublished staff training handbook, 2005.

[PV05] PAQUET E., VIKTOR H.: Long-term preservation of 3D cultural heritage data related to architectural sites. In *Proceedings of the ISPRS WG V/4 Workshop 3D-ARCH 2005* (2005).

[Sav97] SAVEDOFF B.: Escaping reality: Digital imagery and the resources of photography. *The Journal of Aesthetics and Art Criticism 55* (1997), 201–214.

[Sim69] SIMMONDS H.: *Archaeological Photography*. University of London Press, 1969.

[Tec02] TECHNICAL ADVISORY SERVICE FOR IMAGE: *An Introduction to Digital Preservation*, 2002. URL = http://www.tasi.ac.uk/advice/delivering/digpres.html, accessed 20/04/06.

[Woo04] WOOLLISCROFT D.: *Thoughts on the suitability of digital photography for archaeological recording*. Institute of Field Archaeologists Opinion Page, 2004.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

New Reflection Transformation Imaging Methods for Rock Art and Multiple-Viewpoint Display

Mark Mudge,[+] Tom Malzbender,[*] Carla Schroer,[+] Marlin Lum[+]

[+]Cultural Heritage Imaging, San Francisco, California USA
[*] Hewlett-Packard Laboratories, Palo Alto, California USA

Abstract

We offer two new methods of documenting and communicating cultural heritage information using Reflection Transformation Imaging (RTI). One imaging method is able to acquire Polynomial Texture Maps (PTMs) of 3D rock art possessing a large range of sizes, shapes, and environmental contexts. Unlike existing PTM capture methods requiring known light source positions, we rely on the user to position a handheld light source, and recover the lighting direction from the specular highlights produced on a black sphere included in the field of view captured by the camera. The acquisition method is simple, fast, very low cost, and easy to learn. A complementary method of integrating digital RTI representations of subjects from multiple viewpoints is also presented. It permits RTI examination "in the round" in a unified, interactive, image-based representation. Collaborative tests between Cultural Heritage Imaging, Hewlett-Packard Labs, and the UNESCO Prehistoric Rock-Art Sites in the Côa Valley, a World Heritage Site in Portugal, suggest this approach will be very beneficial when applied to paleolithic petroglyphs of various sizes, both in the field and in the laboratory. These benefits over current standards of best practice can be generalized to a broad range of cultural heritage material.

1. Introduction

Paleolithic rock art is among the earliest surviving examples of humanity's cultural heritage (CH). For millennia, the heirs of these ancient rock carvers have added their strokes on stone to our collective legacy. These works are found throughout the world.

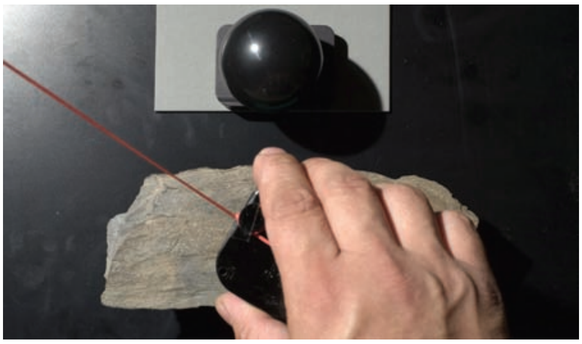


Figure 1: Highlight RTI method

We will first review the state of current CH documentary best practices for rock art and related fields. Next, we will explore test uses of techniques that capture the 3D geometry and textures of rock art. Following a brief description of Reflection Transformation Imaging (RTI) and one of its chief tools, Polynomial Texture Maps (PTMs), we will survey its previous use and advantages in CH contexts with an eye towards potential advantages in rock art. Existing methods of PTM capture using known light positions will be evaluated and approaches for building PTMs without known light positions investigated. These investigations will suggest a new tools and methods set called "Highlight RTI." We will examine how Highlight RTI brings powerful and practical new tools to the study of ancient rock art and how these benefits can be generalized to the fields of epigraphy, sculpture, and other areas of CH work, when materials have significant surface information. We will also see how PTM Object Movies can greatly extend the documentary power, remote ability for interactive scholarly examination, and public access to non-planar CH objects with surface relief.

2. Evaluation of Current Documentary Best Practice in Rock Art and Related Fields.

There is extensive literature regarding the established techniques employed in the documentation of rock art. Despite knowledge of the experimental use of 3D geometry-based imaging methods discussed in section 3, most current CH best practices at the beginning of the 21st-century still rely on the essentially 19th-century methods of hand-drawn graphics and conventional, including digital, photographs. This is especially true of rock art documentation and management.

The limitations of drawings and photographs have been broadly discussed. Among the most significant are the editorial decisions of the draftsperson or photographer to record some elements of the empirical data presented by the subject and exclude others. This is obvious in drawing and is present in photography due to illumination choices in image composition. This problem is understood by epigraphic photographers who decipher inscriptions. They recognize that lighting direction significantly determines the content of the empirical data set recovered. To partially mitigate this problem, they capture multiple images of the subject from the same point of view using different illumination directions to increase the amount of available data [Zuc04].

The incomplete nature of rock art drawing and photography data makes them insufficient to track the effects of natural destructive processes [BCD*05] [FER06] and human contact. Rock art is found in 3D rather than 2D contexts. This 3D context is important to its understanding and is poorly communicated in drawings and photographs.

3. Experimental Acquisition of 3D geometry in Rock Art

Lasers have been used in many CH applications. In the United Kingdom, their use has been demonstrated to benchmark rock art geometry in order to monitor future degradation as part of an integrated management system [BCD*05].

Since 1979 [TWDM79], extensive literature has described numerous applications of photogrammetry to capture 3D geometry in archaeology and rock art around the world. In the United States for example, Federal researchers are looking for simple and inexpensive ways to incorporate these tools into standard documentary practices for rock art and paleontological material on public lands [MN06]. In Europe, photogrammetry-based tools using uncalibrated, handheld photographs and video to generate 3D geometry of CH materials are being explored [MURweb]. A new EPOCH service that accepts a set of such photographs of a single side of a CH object and returns 3D geometry shows promising results [EPOCHweb].

Photometric stereo has been investigated to capture 3D geometry of surface relief materials [EPD04,]. It has also been used, in combination with RTI enhancement, to capture single views of cuneiform tablets [WVM*05].

There are a number of reasons for the slow rate of adoption of these new technologies using 3D geometry in CH. The capture and manipulation of 3D geometry require technologically demanding skill sets outside of those present in most professional rock art and archaeological working cultures. Further, communication and visualization of high-resolution 3D geometry over the Internet is problematic.

4. Use of RTI in Cultural Heritage

The use of image-based RTI documentation avoids many of the problems associated with 3D geometry.

RTI was first developed for use in association with PTMs by Malzbender et al in 2001 [MGW01]. In summary, PTMs are an image-based representation of the appearance of a surface under varying lighting directions. Per-pixel surface normals are extracted from the representation, and can be used for not only changing lighting direction interactively, but also performing enhancements to make surface detail more visible. Transformations of these captured reflectance functions, in particular RTI, have been found to be useful. Unlike photometric stereo, RTI information is communicated without the use of 3D geometry, eliminating 3D geometry's associated requirements and costs.

The use of RTI with PTMs in CH documentation projects, the natural sciences, and law enforcement as well as other Image Based Re-lighting (IBRL) applications has been detailed extensively elsewhere, including [MGW01] [Zuc04] [Mud04] [MVSL05] [MS06] [MO05] [HPLweb] [CHIweb].

This work has demonstrated that RTI with PTMs brings significant advantages to CH activities. These advantages include: non-contact acquisition of data, clear representation of 3D shape characteristics through interactive viewing tools, data discernment improvements over direct physical examination through RTI enhancement functions, no data loss due to shadows and specular highlights, high resolution sample densities up to 20,500 per square millimeter, automatic determination of the most informationally rich illumination directions, simple and achievable image creation processing pipeline, and easy online communication.

In spite of these major advantages over current practice, the bottleneck slowing the adoption of this technology by CH professionals has been existing PTM capture methods. Our new methods described in sections 7 and 8 should remove many of these impediments.

5. Advantages and Limitations of Existing PTM Capture Hardware.

Existing methods of PTM capture all rely on a prior knowledge of the light positions that will be used to illuminate the imaging subject. This knowledge is encapsulated either in a physical equipment structure such as those used by Hewlett-Packard Labs and Cultural Heritage Imaging [MGW01] [MS06] or an instructional template that gives directions for the placement of free-standing lights in pre-determined locations [MVSL05].

The use of fixed-light position equipment has many advantages. Automatic control of lights and camera can acquire a PTM with great speed, frequently between five and fifteen minutes. These efficiencies are valuable when large numbers of objects must be captured. This equipment also can support the optics for controlled wavelength imaging of photonically fragile cultural artifacts [MS06].

Fixed-light position equipment also has attributes which limit its use. The light distance from the subject limits the subject's maximum size. This ratio between light radius and object diameter, approximately 3:1 in existing designs, requires that bigger subjects require proportionately larger light distances. As equipment size increases, light power needs, structural requirements, transport difficulties, and costs increase. In the past, this has limited PTM work to smaller objects. Many cultural heritage materials are found in locations that contain obstructions that prohibit encapsulation by fixed-light hemispheric structures. Further, PTM equipment has been expensive and only available through custom fabrication.

Pre-determined templates or on-site measurement of light positions have the potential to capture larger objects and work around obstructions, but also have limitations. They require time to pre-plan and transfer to the site. Careful positioning of light stands according to the template is also time-consuming. In situations where the floor or ground is irregular such as a natural environment, implementation of a template solution becomes extremely complicated and time-consuming. Locations on the ground and height offsets must be calculated without the benefit of simple planar measurements and may require surveying techniques to adapt the template to the topography.

6. Determination of Light Positions for PTMs

In our method, unlike traditional RTI approaches, the 3D locations of the lights used to construct the RTI need not be known or determined at the time the photographic images are recorded. Our RTI acquisition method manages light intensity fluctuations due to changing light subject distances and captures the necessary information about the light's incident direction.

6.1 Background Considerations

To build an RTI image, the direction of the incident illumination must be known. If the direction is known, and the radius of the light to the subject remains constant, the actual distance value of the radius need not be known. Once the direction of the incident illumination is accounted for, this permits variations in pixel intensity over the subject to be understood as indicating variations in the normal directions and albedo of the underlying surface, not light distance. In past RTI capture equipment and methods, the distance of the light to the subject remained constant by virtue of a prior decision, either in the form of mechanical design or positioning template, constraining the positions of the lights to the geometry of hemispheres or arcs. Without this prior knowledge, if variations in radius are present in the input photographs, they must be managed in order to avoid misinterpreting pixel intensity changes due to changes in light distance as due to the subject's surface features, thereby introducing errors into the normal calculations.

When changes of light source radius are introduced intentionally, these corrections in illumination intensity can be made during the acquisition session by increasing or decreasing the light source brightness at the same time as the radius change. The parameters of these required light power adjustments for distance are well understood and can be looked up in "cheat sheets", covering the distances likely to be encountered. The illumination radius can then be varied for subsets of images during a capture session according to site conditions, for example, shortened to avoid obstructing objects. For example, when the subject is close to the ground and the ground plane acts as an obstruction, adjustments to the illumination radius can be highly advantageous to increase the sampled light directions.

In theory, changes of light source radius could also be recorded automatically at the time of photographic image capture or be extracted later during post-processing. The subsequent illumination variations could then be corrected by adjusting pixel luminance levels

In practice, the limited dynamic range of the camera sensor and the requirement to achieve good exposures of the subject significantly restrict such post-acquisition pixel luminance adjustment options.

The best exposure for a given RTI subject is determined iteratively by checking test images captured at the extremes of the desired range of incident light inclinations. The goal of a successful exposure is to capture as much signal (light) as possible while avoiding under- and over-exposure at these extremes.

As the following histograms from the image set used to build the PTM in Figure 8 indicate, the entire usable exposure range is bracketed by the 2 images at the inclination extremes. In fact, images from the 4 highest inclinations were excluded from the input set due to over-exposure.

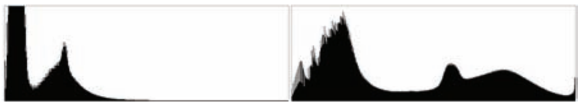


Figure 2: PTM input image histograms from low- and high-incident light angles.

As changes in light source radius lead to geometric increases or deceases of light intensity, the range of light source radius variation, uncorrected by light power adjustment, is very limited without image loss due to improper exposure. This limitation curtails the ability of post-processing methods that adjust pixel luminance levels to successfully correct for these changes. Even when all images are photographed in 12-to-14-bit-per-color channel RAW mode, which has greater dynamic range than in camera compressed 8-bit-per-color channel images, and exposure compensation is performed prior to conversion to 8-bit mode, signal-to-noise ratio concerns reduce the desirability of these operations.

Perhaps most importantly, given that estimations of proper exposure parameters for a given subject and environment are subject to operator error and that the consequences of these errors are potentially costly, reservation of RAW mode exposure compensation margins to permit recovery of these mistakes is highly desirable.

6.2 Evaluation of Light Source Radius Management Techniques

We explored 3 methods of light source radius management; a laser range finder, measurement of highlight intensity from a black ball, and a string marked to known lengths.

A Leica Disto laser range finder was tested during field capture of petroglyphs in Portugal. The Disto was secured to a small ball head mounted on a bar next to the light and could be pointed and locked in the desired orientation. The center of the subject imaged in the camera was identified and a distance reading taken to position the handheld light at the radius appropriate to the illumination power setting.

We found that positioning the light within 10cm of the stipulated 250cm radius took an average of five readings and considerable time. It was also hard to use if the light was held overhead or near the ground. The Disto added weight to the light setup and significant cost.

Capturing the intensity of a black ball highlight for potential use in post processing to compensate for light source radius variations was explored. The contrast ratios between the black ball highlights and normally exposed RTI subjects exceed the single image dynamic range of today's digital cameras. To avoid data loss of the highlight intensity data due to clipping from over-exposure, a 6 stop under-exposure was needed to extract the relative highlight intensity information. This required the use of a 3 image exposure bracketing sequence.

As the PTM process requires multiple input images, we found that the tripling of capture data this process entailed was a serious disadvantage.

A string marked at known distances was also tested. One end of the string is tied to the light source, operated by one person, and the other end is held by another person adjacent to the center of the subject. The light distance is measured, the light held steady, the string moved out of the image area, and the photo taken. The radius can be changed during the session by changing the string length.

The string, nicknamed "The Egyptian Method," became our favored light source radius management method. The string has many comparative advantages. It is intuitive and easy to learn. It is fast. Once camera configuration is complete, 60 PTM photographs can be captured in 20 minutes. With a handy chart of distance-dependent light power values, changing the string length and light power at the same time is simple and easy to remember.

6.3 Previous Image-Based Relighting Work

The use of spheres to collect lighting direction of a handheld light source for the purposes of image based relighting was first introduced by Vincent Masselus et al. in [MDA02]. In Masseulus's approach, 4 diffuse white spheres are placed in the field of view surrounding the object in question. Two synchronized cameras are used, one which is focused on the object being photographed, and the second, with a larger field of view, that captures images of the spheres as well. Since the spheres themselves are diffuse, a slowly varying Lambertian model is fit to the sphere surface to recover the lighting direction.

In our approach we use a single highly specular black sphere which returns a high-contrast highlight that is very easy to detect and localize. In addition, we employ only a single camera that captures both the highlight location (and thus light direction), as well as the surface being studied.

More similar to our approach is that of Paul Debevec's group in reconstructing archaeological inscriptions, which is briefly described in [EPD04]. Two glossy black spheres and the subject are imaged in a single photograph, light vectors are extracted from highlight positions and intersected.

7. Lighting Direction from Highlight Locations

To avoid using an elaborate light stage [DHT*00] [MGW01] with known light source positions, we rely on the user to position a handheld light source, and recover the lighting direction from the specular highlights produced on a black sphere included in the field of view captured by the camera. Since the specular highlights can be of higher brightness than that of the rest of the illuminated scene, we need to be careful to limit the amount of saturation so as not to hamper our estimation of the center of the highlight location. Currently we allow some of these highlight pixels to be saturated, giving priority to the exposure of the scene itself. Manually estimating the position of the center of each highlight is still practical in this circumstance. In the future we hope to be able to automatically detect the sphere in all the images, and extract its radius along with the location of the center of the highlight automatically. Note that since the camera is stationary for every photograph, the location of the sphere remains constant, and this appears feasible to extract.

Given a specific photograph we solve for the unknown light direction, L, referring to Figure 3. We assume the distance to the camera and lights is large with respect to the object geometry to simplify the solution. First we measure (in pixel coordinates) the center (C_x,C_y) and radius, r, of the sphere and the position of each highlight center (H_x,H_y). We then normalize the highlight location to span (-1,1) by:

$$S_x = (H_x - C_x) / r$$

$$S_y = (H_y - C_y) / r$$

Thus the normal vector is given by:

$$(S_x, S_y, \sqrt{1 - S_x^2 - S_y^2})$$

Performing a Cartesian to polar conversion, the inclination of the surface normal, Φ, is seen as:

$$\Phi = \cos^{-1}(\sqrt{1 - S_x^2 - S_y^2})$$

Since the inclination of the light source, Φ_L, is simply the inclination of the normal Φ, it is just:

$$\Phi_L = 2\cos^{-1}(\sqrt{1 - S_x^{\,2} - S_y^{\,2}})$$

For y>0, the other angle of the light source in spherical coordinates is seen as:

$$\Theta_L = \sin^{-1}(y/\sin(\Phi))$$

So, the normalized light vector, L, is given by:

$$L_x = \sin(\Phi_L)\cos(\Theta_L)$$

$$L_x = \sin(\Phi_L)\sin(\Theta_L)$$

$$L_z = \cos(\Phi_L)$$

These light vectors, one for each photograph taken, can now be used to construct a PTM, allowing subsequent enhancements.

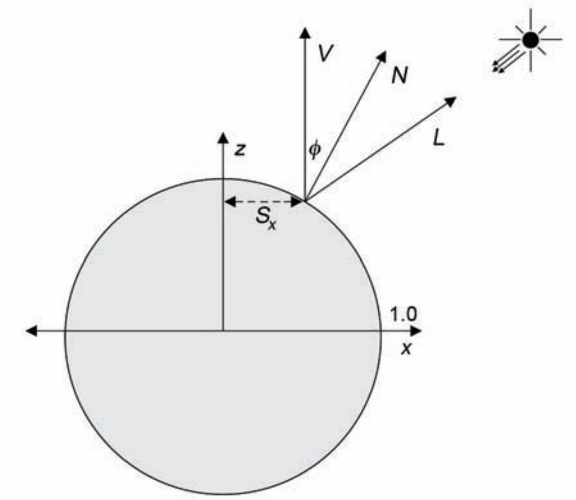


Figure 3: Estimating lighting direction from a highlight location on the surface of a glossy sphere assuming a distant camera. V is the view vector pointing to the camera, N is the surface normal, L is the unknown, normalized light vector we solve for.

8. Work Tools and Methods

The field kit contains: a SLR digital camera; a configurable selection of light-reducing neutral density filters; computer with remote camera control software and Photoshop for evaluating images; tripods; radio flash trigger; 1 to 1/32 power adjustable 320 watt second flash with battery pack; black ball on length and angle adjustable boom; measuring tape; and retractable surveyors plumb-bob string (bob removed). An intensity-adjustable continuous light may also be used if sufficient power is available.

Ambient light is excluded using the neutral density filters. Some ambient light is excluded by a narrow aperture, usually f/11 to f/16, dictated by depth of field and image sharpness considerations and relatively fast shutter speed for flash synchronization, usually 1/180th of a second. The filters are added in the strength necessary to remove any significant levels of remaining light. Subsequently, the only light reaching the sensor during image acquisition comes from the flash. The light source used was an inexpensive, mid-power, studio light costing $300. It was sufficient to provide more than enough light, even shooting in full sun at midday. The blackball is placed next to, but not touching, the subject. The image is composed to maximize the number of available pixels on the subject. The operators capture test images verifying ambient light exclusion, focus, depth of field, and exposure. Images are captured in RAW format. A data set of at least 60 images is collected using different light locations. The selection is biased towards the first 60 degrees of inclination from the "horizon."

Figure 4: Field setup

During image processing, any lens distortion is corrected. The black ball's pixel radius, and center in (X,Y) pixel coordinates are determined. The light's highlight center coordinates on the ball in each image are collected. This information is placed in a text file that is given to the Highlight to Light Position (HLT-LP) software and a Light Position ".lp" file is generated. Remaining processing is done according to standard PTM processing procedures.

At the end of the process, an archive is created for each PTM, including RAW data, Photoshop actions, xmp files, HLT files, .lp files, PTMFitter command syntax, and descriptive text enabling interested parties to review the image processing steps employed to build it. This information can be used for confirmatory replication. In this way, the quality of each PTM image can be independently assessed.

9. PTM Object Movies

PTM Object Movies (POMs) synthesize individual PTMs into an integrated, interactive representation.

9.1 The PTM Object Movie Viewer

The PTM object movie (POM) viewer software takes as input a text file indicating the number of individual PTM images used per inclination angle row, the number of rows, and their identifying absolute pathnames. It then assembles them using established object movie creation operations and displays them in our new viewer. Unlike traditional object movies, when a specific point of view is selected, the user may employ a menu of standard RTI enhancement functions to disclose the subject's 3D surface features.

9.2 PTM Object Movie Tools and Methods

For the single row POM presented here, we used the automatic PTM acquisition structure in Figure 5 [MS06]. This structure also possesses a detachable module, not shown, that permits manual vertical inclination adjustment. The subject was rotated around the camera image's Y axis using a programmable electronic object movie/panorama stage from Seitz Phototechnic. The stage platform was custom built to create a top face edge profile with size and shape characteristics identical to the subject's size and shape. This prevents shadowing of the subject by the platform when the incident light comes from below the platform.

Figure 5: POM setup

The tools and methods of acquiring input images closely resemble the procedures for making normal single and multi-row object movies. As in the traditional technique, the subject is sampled by capturing individual PTMs at regular rotational intervals while maintaining a constant camera-to-subject radius. While light position imaging is introduced into the process, much of the process and equipment design strategy used in normal object movie camera location is applicable.

The highlight PTM equipment may also be used to capture the PTMs used to build a POM. As with normal object movies, when using the Highlight PTM technique, the camera positions used to photograph the input PTMs require management to ensure constant viewpoint sampling increments and camera-to-subject distance. This can be accomplished using a predetermined template or calculating the correct camera position during capture and is subject to the time and complexity issues described previously. The precision needed in these calculations is determined by focus depth of field, and the user's aesthetic tolerance of subject size and orientation irregularities in the POM.

10. Results

With the help of our Portuguese collaborators at the Parque Arqueologico do Vale do Côa (PAVC), Centro Nacional de Arte Rupestre (CNART), and the Universidade do Minho, we captured five PTMs reflecting a sample of the different types of rock art at the site.

Figure 6: PTM image of Penascosa Roche 5B, depicting paleolithic animation techniques.

The subject of Figure 6 was created during the Magdalenian period. It shows a left-facing goat profile with the head and neck 'animated' to indicate both forward- and backward-looking orientations [Bap99].

The data set images captured a 68cm by 46cm area on the rock face. Given the 4096-by-2704 pixel image resolution, 3D surface normal information was recorded every 166 microns or 36 samples per square millimeter. The final image was cropped to represent an area 45cm square. The session lasted 105 minutes, including equipment set up, tear down, and image testing. Photography of the 60 varying light direction images in the set took 30 minutes. The session took place between 12:00 noon and 1:45pm on a sunny, cloudless day. The PTM clearly demonstrates the success of ambient light exclusion using neutral density filter combinations.

Figure 7 shows two areas selected from Figure 6. The specular enhancement applied to Figure 7 shows the temporal sequence of two intersecting engraved lines. The diagonal line describing the dorsal profile of the forward facing goat head was clearly cut through the more horizontal line. This establishes the goat engraving as younger than the horizontal line engraving.

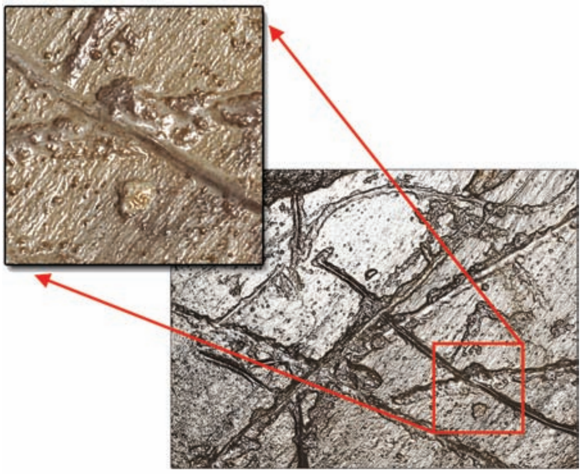


Figure 7: The engraving sequence

The subject of Figure 8 is 18cm wide and the PTM sampled 3D normal information every 65.7 microns or 231 times per square millimeter. The tiny petroglyph featured is only 3.1cm in length and hard to see, both in photographs and during direct physical examination.

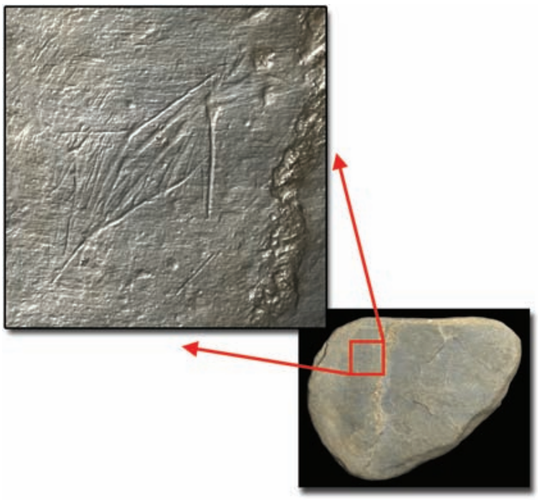


Figure 8: A very small and fine drawing from a stone used as portable art. It was found during an archaeological excavation at the PAVC in 1999 and was in the context of other evidence of Magdalenian human habitation.

In contrast, RTI imaging functions permit discernment of the most difficult to see surface attributes. The petroglyph-rich front and back of the stone and the engraved face of a second similar stone were imaged during a single day by the professional staff from Portuguese Centro Nacional de Arte Rupestre. After set up and imaging tests, the PTM data capture of 60 images for each of the 3 sets averaged 20 minutes. Under our initial supervision, the local team conducted the entire capture session for the 3 data sets after only a half day of instruction in the Highlight PTM method. As we gave them a black ball and their existing inventory of photographic equipment included everything necessary to capture Highlight PTMs, there was no equipment start up cost.

Figure 9 shows different views of the subject as seen in the POM Viewer. The subject is a Bronze Age Torque in the collection of the Congregation of the Grand St. Bernard in Switzerland. It was worn around the neck of a person buried sometime between the 20th and 16th centuries BCE. It displays fine and intricate engraving over its entire exterior surface. This continuous abstract geometric design can be followed all around the circumference of the torque and examined at any position using RTI enhancement.

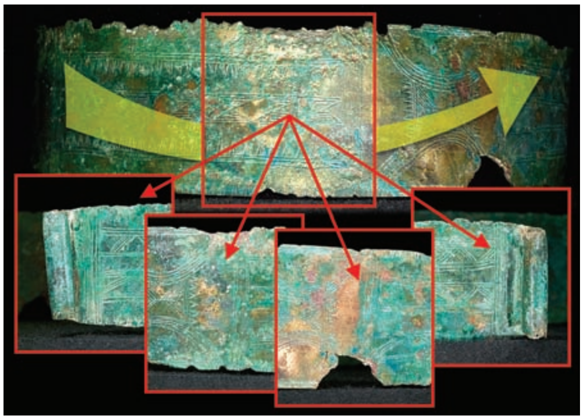


Figure 9: Multiple RTI views of an object

The subject can be interactively rotated in the POM viewer and the lighting changed as desired by the user.

11. Conclusions and Future Work

The Highlight PTM method is able to acquire Polynomial Texture Maps (PTMs) of 3D rock art possessing a large range of sizes, shapes, and environmental contexts. The acquisition method is simple, fast, very low cost, and easy to learn. Our experiences working with our collaborators at CNART proved that learning the Highlight PTM technique takes less than a day. The simple, intuitive character of the method focuses the attention of the acquisition team on getting the proper photographic exposure at the time of capture.

The POM tools permit RTI examination of CH material “in the round’” in a unified, interactive, image-based representation. POMs offer a powerful new approach to virtually communicate the complete 3D shape and surface information of CH materials without reliance on full reconstruction of 3D geometry and associated processes.

Future work includes fully automating Highlight PTM post processing and mapping the full PTM and POM image generation process history to the CIDOC Conceptual

Reference Model. Additional functions will be added to PTM viewing tools that will more fully adapt them to the normal workflow of the CH professional that will use them. These functions will include the ability to make drawings on top of PTM images while panning and zooming. They will also feature the ability to output RTI-enhanced images and image sequences in a variety of formats.

12. Acknowledgements

We wish to thank our collaborators: Antonio Martinho Baptista, Alexandra Cerveira Lima, Luis Paolo Santos, Alberto Proenca, Friar Luis Proenca, Manuel Almeida, Luis Luis, Andre Tomas Santos and Chanoine Jean-Pierre Voutaz. Thanks to Hewlett-Packard Labs and the Cultural Herititage Imaging Board of Directors, contributors, and volunteers for support of this work.

13. References

[Bap99] Baptista, A. M., 1999. *No tempo sem tempo: A arte dos cacadores paleoliticos do Vale do Côa*, Centro Nacional de Arte Rupestre e Autor. ISBN: 972-98121-0-1.

[BCD*05] Barnett, T., Chalmers A., Díaz-Andreu M.,. Ellis G, Longhurst P., Sharpe K., Trinks I., 2005. 3D Laser Scanning For Recording and Monitoring Rock Art Erosion, *International Newsletter on Rock Art (INORA)* 41: 25-29.

[CHIweb1]http://c-h-i.org/examples/ptm/ptm.html

[DHT*00] Debevec, P., Hawkins, T., Tchou, C., Duiker, H., Sarokin, W., Sagar, M., 2000. Acquiring the Reflectance Field of a Human Face, *Proceedings of Siggraph 2000*, pp. 145-156.

[EPD04] Einarsson, P., Hawkins, T., Debevec, P., 2004. Photometric Stereo for Archeological Inscriptions, *Siggraph Sketch*.

[EPOCHweb] EPOCH 3D Webservice. http://homes.esat.kuleuven.be/~visit3d/webservice/html/

[Fer06] Fernandez, A. P. B., 2006. Understanding an Unique Conservation Work Environment: The Case of the Côa Valley Rock Art Outcrops. *Theory and Practice in Conservation: A Tribute do Cesare Brandi (Proceedings of the International Seminar)*, Rodrigues, J. D., Mimoso, J. M., eds. Lisboa: Laboratório Nacional de Engenharia Civil, p. 323-332.

[HPLweb] http://www.hpl.hp.com/research/ptm/

[MDA02] Masselus, V., Dutre, P., Anrys, F., 2002. ”The Free-form Light Stage”, *Eurographics Workshop on Rendering*, Pisa, Italy, June 26-28, 2002.

[MGW01] Malzbender, T., Gelb, D., and Wolters, H. 2001. Polynomial texture maps**.** *Proceedings of ACM Siggraph 2001*.

[MN06] Mathews, N., Noble, T., 2006. Surveying by Photo: Get Close-range Photogrammetric Control that's Accurate Enough. Paper delivered at the *Computer Applications and Quantitative Methods in Archaeology Conference* (CAA2006), Fargo, ND USA.

[MO05] Malzbender, T., Ordentlict, E. Maximum Entropy Lighting for Physical Objects. *Hewlett-Packard Technical Report* HPL-2005-68. http://www.hpl.hp.com/personal/Tom_Malzbender/papers/papers.htm

[MS06] Mudge, M. Schroer, C., 2006. 3D Data Acquisition with Reflection Transformation Imaging. Paper delivered at the *Computer Applications and Quantitative Methods in Archaeology Conference* (CAA2006), Fargo, ND USA.

[Mud04] Mudge, M., 2004. Web Graphics/Special Sessions/Panels, Cultural Heritage and Computer Graphics Panel, *SIGGRAPH 2004 Conference Presentations*, Soma Media, publisher, ISBN1-58113-950-X.

[MURweb] 3D Murale project website. http://dea.brunel.ac.uk/project/murale/murale.html

[MVSL05] Mudge M., Voutaz J.P., Schroer C., Lum M., 2005. Reflection Transformation Imaging and Virtual Representations of Coins from the Hospice of the Grand St. Bernard. In *Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2005)*, Mudge M., Ryan N., Scopigno R. (Eds.), Eurographics Association, pp. 29– 39.

[TWDM79] Turpin, S. A., Watson, R. P., Dennett, S., and Muessig, H., 1979. Stereophotogrammetric Documentation of Exposed Archaeological Features in *Journal of Field Archaeology.*

[WVM*05] Willems, G., Verbiest, F., Moreau, W., Hameeuw, H., Van Lerberghe, K., Van Gool, L., 2005. Easy and cost-effective cuneiform digitizing. In *Short and Project Papers Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2005)*, Mudge M., Ryan N., Scopigno R., (Eds.), Eurographics Association, pp. 73-80.

[Zuc04] Zuckerman, B., 2004. Every dot and tiddle: A consideration of the limitations of computer imaging for the study of Dead Sea scrolls. *Double Takes: Thinking and Rethinking Issues of Modern Judaism in Ancient Contexts* Garber, Z. and Zuckerman, B. (Eds) Studies of the Shoah Volume XXVI, University Press of America

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

An Ontology for 3D Cultural Objects

Franco Niccolucci[1] and Andrea D'Andrea[2]

[1]PIN, Prato, Italy
[2]CISA, University of Naples "L'Orientale", Naples, Italy

Abstract

3D cultural objects are digital 3D replicas of objects having a cultural value, as models of artefacts, reconstructions of buildings, sites and landscapes. As such, they have a twofold nature, and inherit properties both from their digital nature, like the shape and texture, and from the cultural content, for instance to be used for scholarly purposes or communication to the public. In some cases, one of the natures prevails on the other. This may be the case because the object is being processed, e.g. visualized on a computer, or scrutinized by heritage scholars for review. In a few others, it is unfortunately the user's narrow-minded attitude that leads to take into account only one nature of such an object and neglect the other. It is therefore necessary to explore a way of documenting 3D cultural objects that keeps together all the relevant information, both the cultural and the digital one. In this paper we propose an ontology for such complex objects that owns the following important properties: i) it is sufficiently general to encompass very different artefacts, from pottery sherds to historical landscapes; ii) it fully complies with international standards for heritage, in this case CIDOC-CRM, of which it can be shown to be a specialization/extension; iii) it is sufficiently simple to be used and understood by heritage practitioners and professionals with moderate computer skills, and documents items in a plain, human readable and understandable way; iv) items documented as instances of this ontology can be efficiently processed for the most frequent purposes, as computer visualization, retrieval of cultural information or storage in a database; v) it is ready for compliance with other important requirements, as for instance the proposed charter on credibility known as London Charter.

Categories and Subject Descriptors (according to ACM CCS): H.3.7 [Standards]:

1. Introduction

A holistic approach to computer visual representations of heritage artefacts is being fostered by a number of scholars investigating not only how Information and Communication Technology (ICT) may enhance the visualization of such objects, adding spectacular features to cultural communication, but also the rules that modelling and visualization must follow to respect the cultural content. Such an approach is the foundation of the EU project EPOCH [EPO], which is now developing an interoperable set of tools by combining existing ones into a toolkit and a common software infrastructure. However, it is evident that such efforts requires the adoption of standards at the interfaces between different tools and at the beginning and end of the production chain. Therefore, standards for 3D objects are currently investigated, in order to optimize technology and guarantee the cultural validity of outcomes. These two facets of cultural objects must march together, indissolubly tied also as far as data are concerned. In both regards standards do exist.

Technology offers a variety of *de facto* or *de jure* possibilities, presently being investigated by the project, while for the cultural side CIDOC-CRM [CRM] is the necessary choice. As it is well known, CICOC-CRM is an International ISO Standard (ISO 21127:2006) under publication as of 06/06/06. It "establishes guidelines for the exchange of information between cultural heritage institutions. In simple terms this can be defined as the curated knowledge of museums. [...] ISO 21127:2006 is specifically intended to cover contextual information (i.e. the historical, geographical, and theoretical background that gives museum collections much of their cultural significance and value)." [ISO06]. CIDOC-CRM current version is 4.2.

At this point, there is the need of a container to host the different descriptors of the 3D cultural object, the digital replica of a cultural artefact reproducing a man-made object or a complex of such objects and natural ones. The definition of "real" cultural objects, as compared to "digital" cultural objects, is beyond the limits of the present paper. However, it assumed that CIDOC-CRM is a "good" way to describe them, i.e. it satisfactorily manages the necessities of documentation of the cultural heritage domain. Therefore we will assume that the objects originating the digital replicas we are going to consider here are (well) described using CIDOC-CRM. Calling them cultural objects is just shorthand and it just implicitly means that whatever the definition of cultural heritage objects one has in mind, CIDOC-CRM is an ontology describing the things that belong to it.

It has been recently noted that just documenting the cultural object and the associated digital object is not sufficient to guarantee credibility to the latter. Draft guidelines to document such derivation process, the so-called "London Charter", are in progress [BDN06]. In our opinion, this documentation is the third facet of the digital cultural object description, and is the glue that keeps together the other two. In fact, many digital objects (DO) may be associated to the same cultural object (CO) using different replication processes (RP), as scanning, modelling and so on, and only triplets (CO, DO, RP) are unique: RP: CO $\mapsto$ *DO*. By the way, unless random factors are deliberately introduced — in this context, in a rather unjustified way — the description of RP guarantees in principle that the derivation may be repeated. In practice this may be limited by external factors such as, for example, environment illumination in scanning, although this is true for any scientific experiment and repeatability is generally only theoretical.

Another factor adds to the difficulty of keeping the above descriptions together by attempting to reconcile the respective different ontologies. According to [HPS04], "a reason for expecting continuing diversity is that an ontology is often aligned with a particular perspective on the world." In our case, there are two perspectives, the cultural one (which cares little of the way the object is digitalized, as long as some validity rules are respected) and the technological one (which pays attention to the technology but needs guidance as far as embedded "cultural" content is concerned). Such perspectives correspond to two different communities, heritage professionals and computer experts; reconciliation comes from the necessity of interacting on a common goal, for example communicating heritage availing of computer graphics.

In fact, both parties often state that the "other" feature is beyond their scope. For instance, [Doe01] describes a comparison of OpenGis with CIDOC-CRM and, when technical details are involved, rightfully states that "this is out of the scope of the CRM." In a similar way, technical 3D standards as X3D [X3Da] or Collada [Col] have little room, if any, for incorporating information concerning the"original" being modelled.

Therefore, the optimal solution would be to enable each party to keep their own ontology and related perspective of the world, and to provide tools to combine the different ontologies into a more general one, that can be "flattened" back into generalisations of the original ontologies, minimally extended with a container for off-topic information; a simpler task, if the original ontologies already have a container for it. In other words, the mechanism we propose is such to allow each party to deal with familiar concepts, with perhaps just a small extension, where all the non-pertinent stuff is stored: that is, heritage professionals are enabled to manage information with CIDOC-CRM and accept that there is a class where the digital information is stored; technology professionals will have a place in the standard they use for storing cultural information which may be irrelevant for processing.

The difficulty of this approach is that it must be provided a mechanism for collapsing the information which is irrelevant for a particular perspective while expanding the one which is, on the contrary, relevant, from a background where both pacifically coexist.

This is what we will try to do in the present paper.

Since the methodology for the documentation of the production process is still in its infancy, we will limit to the other two facets, i.e. we will consider the ontology of Cultural Objects (CIDOC-CRM), which concerns the CO; and the ontology of 3D objects (in particular, X3D, for reasons to be clarified below), which concerns the DO, considering only very simple information relating to the RP to be stored wherever appropriate.

2. Previous work: combining the CRM with MPEG

Previous work in this direction has been undertaken by Jane Hunter in [Hun02]. In this paper the author describes an approach which combines the domain-specific aspects of MPEG-7 and CIDOC-CRM into a single ontology, in order to describe and manage multimedia in museums. She also gives the complete description of an example concerning an ethnographic film taken between 1901 and 1912, and digitized in 1999 for preservation purposes.

Although listing a number of Museum Multimedia Types, the author's focus is video-audio documentation of intangible heritage. In this case the digital object is not a replica, but more precisely a substitute. Incidentally, also in this case a description of the production process would have been substantial: did the movie film aboriginal dances taking place independently, or the director arranged people to perform dancing in front of the camera? Did the digital copy include all of the original movie, or some parts were lost?

Hunter's paper then describes a combined ontology in which CIDOC-CRM is extended to include MPEG-7 classes, attached to CIDC-CRM where they fit better.

Some information is present in both ontologies, and may be directly mapped. Other is better managed by one of the two merging ontologies. The attachment point for the MPEG-7 class hierarchy is CIDOC-CRM class E73.InformationObject, which is enriched by adding appropriate subclasses.

Although paving the way for the model we will discuss in this paper, and being a very good application of the capacity of CIDOC-CRM to manage unforeseen conditions, Hunter's approach cannot satisfactorily manage the problems we are dealing with here. Her examples are in fact the physical carrier of cultural information which is stored nowhere else and are autonomous; digital replicas are on the contrary cultural nonsense, or just computer graphics exercises, out of the heritage context in which they originated.

3. Is a bell beaker a Bell Beaker?

To further proceed in the definition of a harmonized ontology for 3D objects, we will consider a bell beaker like the one reproduced in figure 1. These objects are typical of the Bell Beaker culture, an archaeological culture spread in Europe and dating from 2800 to 1900 BC. In other terms, these objects are associated with a complex of time periods, regions and practices. They are supposed to have been used for drinking beer, or mead, possibly in ceremonies, and are usually part of a prestige package in grave goods.

The one depicted below comes from the Wiltshire Heritage Museum in UK. It is a European Bell Beaker dating before 2500 BC found complete and intact in West Kennet Long Barrow near Avebury. This is perhaps the best-known beaker from the British Isles.

Figure 1: *A Bell Beaker from the Wiltshire Museum.*

Now, assuming the above (or, better, the file from which the image originated) is a faithful reproduction of the Bell Beaker kept in the museum, what makes it a digital cultural object? It is in fact the statement given above, relating it to the real Bell Beaker of the museum. In other words, it is this relation that turns a computer file — the bell beaker — into a digital cultural object — the Bell Beaker digital replica. Is the latter a cultural object by itself? In general no, unless we want to consider it as artwork, what is certainly not the case for Figure 1. This might the case, for instance, of the well-known Picasso's painting *Las Meninas*, a set of variations after the homonymous painting by Velazquez. So, using an ontology as CIDOC-CRM, conceived for cultural heritage, to document it might appear inappropriate, and an overarching technological standard as MPEG-21 might seem to work better. This is not the case, unless there is a way for embedding in it the cultural information in a simple and elegant way — and current standards do not envisage such opportunity. A possibility would be to extend existing technological standards, for example creating a "cultural profile" in X3D with appropriate extensions as suggested in [Nic06]. Although possible, this would create yet another "standard" of difficult maintenance. Even if inspired by an accepted standard as CIDOC-CRM, who would guarantee that any future change in the latter would be implemented in such extensions? So the only way to proceed is to incorporate into CIDOC-CRM extensions taken from technological standards, but providing automatic mechanisms for update when future versions of the technological standards are made available.

4. Not extending, but embedding

The process of extension is sometimes considered as the addition of new features to an existing system. This is in fact shorthand for a more complex process:

1. Start from an ontology O
2. Define a new ontology O_1 which has a subset O_2 which is isomorphic to O
3. Identify O_2 with O.

If the process is static, the two ontologies O and O_1 will diverge after extension. New versions of O will likely be no more isomorphic to O_2. If the process is dynamic, any change on O will reflect on O_1 in such a way that O_2 will continue to be isomorphic with the new O.

The Web Ontology language OWL [OWL04] provides such a mechanism for dynamic extension. It is the include-style owl:imports construct that together with namespace declaration allows using concepts (classes) from any existing ontology. Appropriate usage of the property owl:equivalentClass eliminates duplicates.

A (partial) description of CIDOC-CRM in OWL has been given in [Bal05] basing on previous work in RDFS and DAML+OIL . It refers to version 3.4.9 and although accessible over the Internet at the designated URI, it is not likely to be maintained, or at least subsequent version will be possibly stored elsewhere, according to the convention apparently used for the URI. An URI system like the one used by W3C, storing the latest version always in the same

URI would better help updating imports. Also, a more significant name than "index.html" would be more appropriate, and perhaps an ".rdfs" file extension would help managing it with ontology tools. However, it suffices for the goals of the present paper.

The EU project DELOS, an IST Network of Excellence on Digital Libraries, has developed an OWL description of the X3D ISO standard [CK06] named OntologyX3D. Although possible for other standards, we will develop the full construction for X3D only, because it is an ISO Standard and has an ontology description, whereas Collada, a potential competitor, has no such facility. For the sake of simplicity in explanation, we will add a universal class named Scene (corresponding to the root element of the X3D Schema) and assume that the ontology is modified accordingly. This addition is perfectly legal and in fact it can be done with no adverse consequence.

So the bricks for building the 3D Cultural Objects Ontology are there, and it remains only to build it up.

5. Building a simple 3D-CO Ontology

5.1. Embedding the parent ontologies

The initial step consist in the definition of the new ontology as deriving from the two parent ones. Firstly, let us define some entities to store the URIs of the parent ontologies:

```
<!DOCTYPE rdf:RDF [
 <!ENTITY cidoc
  "http://cidoc.ics.forth.gr/OWL/crm3.4.9#">
 <!ENTITY ontox3d
  "http://sargos.ced.tuc.gr/OntologyX3D#">
 <!ENTITY owl
  "http://www.w3.org/2002/07/owl#">
 <!ENTITY rdf
  "http://www.w3.org/1999/02/22-rdf-syntax-
ns#">
 <!ENTITY rdfs
  "http://www.w3.org/2000/01/rdf-schema#">
 <!ENTITY xsd
  "http://www.w3.org/2001/XMLSchema#">
 <!ENTITY co
  "http://www.3d-co.org/current#">
]>
```

The first two entities `&cidoc;` and `&ontox3d;` provide shorthand for the respective URIs. Other entities have been added to the DTD to facilitate reading the code in the sequel; they are the standard ones for OWL, RDF, RDFS and XML Schemas. The last entity is a dummy reference to an hypothetical resource (actually *non-existent*) destined to host the 3D-CO ontology under definition.

Then let us include all the relevant namespaces in the OWL header:

```
<rdf:RDF
  xmlns     = "&co;"
  xml:base  = "&co;"
  xmlns:cidoc = "&cidoc;"
  xmlns:ontox3d = "&ontox3d;"
  xmlns:owl = "&owl;"
  xmlns:rdf = "&rdf;"
  xmlns:rdfs = "&rdfs;"
  xmlns:xsd = "&xsd;"
>
```

Now it remains only to import into the 3D-CO ontology the two parent ones:

```
<owl:Ontology rdf:about="">
 <rdfs:comment>
    This is the Ontology
    for 3D Cultural Objects,
    to be used to document
    3D replicas of objects
    with cultural value
 </rdfs:comment>
 <owl:imports rdf:resource="&cidoc"/>
 <owl:imports rdf:resource="&ontox3d"/>
 <rdfs:label>
    3D Cultural Object Ontology
 </rdfs:label>
</owl:Ontology>
```

So far, so good. Now it remains the second part of the task, that is identifying common features and attaching the universal class of OntologyX3D somewhere into CIDOC-CRM

5.2. Identifying common features

The CIDOC-CRM entity `E73.Information Object` is the attachment point for the OntologyX3D superclass Scene. However, since the latter is a particular type of Information Object it probably fits better with the subclass `E36.Visual_Item`:

```
<owl:Class rdf:about="#Model3D">
  <owl:equivalentClass
   rdf:resource="&cidoc;
    E36.Visual_Item"/>
  <owl:equivalentClass
   Rdf:resource="ontox3d;
    Scene"/>
</owl:Class>
```

The above equivalence, established towards CIDOC-CRM classes, projects onto properties, inducing the creation of properties equivalent to those in CIDOC-CRM that relate the visualization to its subject. So `P138.represents` has an homonymous equivalent in 3D-CO, which can be defined within 3D-CO and then associated to `P138` using `owl:equivalentProperty`, or simply quoted as `&cidoc;P138F.represent`.

Here is the resulting ontology as yet obtained.

```
<!DOCTYPE rdf:RDF [
 <!ENTITY cidoc
```

```
  "http://cidoc.ics.forth.gr/OWL/crm3.4.9#">
 <!ENTITY ontox3d
  "http://sargos.ced.tuc.gr/OntologyX3D#">
 <!ENTITY owl
  "http://www.w3.org/2002/07/owl#">
 <!ENTITY rdf
  "http://www.w3.org/1999/02/22-rdf-syntax-
ns#">
 <!ENTITY rdfs
  "http://www.w3.org/2000/01/rdf-schema#">
 <!ENTITY xsd
  "http://www.w3.org/2001/XMLSchema#">
 <!ENTITY co
  "http://www.3d-co.org/current#">
 ]>
<rdf:RDF
  xmlns     = "&co;"
  xml:base  = "&co;"
  xmlns:cidoc = "&cidoc;"
  xmlns:ontox3d = "&ontox3d;"
  xmlns:owl = "&owl;"
  xmlns:rdf = "&rdf;"
  xmlns:rdfs = "&rdfs;"
  xmlns:xsd = "&xsd;"
>
<owl:Ontology rdf:about="">
 <rdfs:comment>
    This is the Ontology
    for 3D Cultural Objects,
    to be used to document
    3D replicas of objects
    with cultural value
 </rdfs:comment>
 <owl:imports rdf:resource="&cidoc"/>
 <owl:imports rdf:resource="&ontox3d"/>
 <rdfs:label>
    3D Cultural Object Ontology
 </rdfs:label>
</owl:Ontology>
<owl:Class rdf:about="#Model3D">
  <owl:equivalentClass
   rdf:resource="&cidoc;
    E36.Visual_Item"/>
  <owl:equivalentClass
   Rdf:resource="ontox3d;
    Scene"/>
</owl:Class>
</rdf:RDF>
```

One might also wish to be more precise and extend CIDOC-CRM with a new class `3D Model`, at the same hierarchy level of `E37.Mark` and `E38.Image` and, like them, a subclass of `E36.Visual_Item`. This option does not alter the substance of our discourse. It can, on the contrary, improve our results. To this goal, let us introduce a subclass of the new ontology, subclass of VisualItem, named Model3D to avoid issues arising with a number as first letter. In this case, it is `co:VisualItem` which is equivalent to `cidoc:E36.Visual_Item`, and X3D is attached one level below — still to `Model3D`, but this has been moved to a lower hierarchy level.

The excerpt of the OWL description is therefore the following:

```
<owl:Class rdf:about="#VisualItem">
  <owl:equivalentClass
   rdf:resource="&cidoc;
    E36.Visual_Item"/>
</owl:Class>
<owl:Class rdf:about="#Model3D">
<rdfs:subClassOf rdf:resource =
     "#VisualItem"/>
  <owl:equivalentClass
   Rdf:resource="ontox3d;
    Scene"/>
</owl:Class>
```

Figure 2 illustrates the above merging process. The leftmost box represents the 3D-CO:VisualItem class, which is equivalent to CIDOC-CRM:E36. The mapping brings into 3D-CO all other CIDOC-CRM classes, including, for example, `E37.Mark` and `E38.Image`. In 3D-CO we define a new class, `3D-CO:Model3D` as a subclass of `3D-CO:VisualItem`. The existing one-to-one correspondence of a subset of 3D-CO and CIDOC-CRM is so preserved. The new class `3D-CO:Model3D` is defined as equivalent to `Scene`, the root of X3D, i.e. `OntologyX3D:Scene`. Further `OntologyX3D` classes are mapped into 3D-CO as subclasses of Model3D.

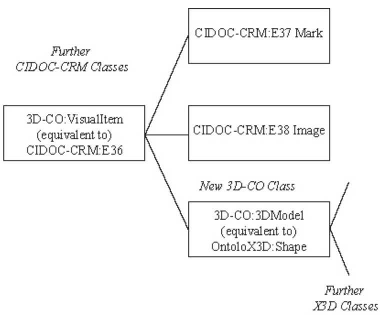


Figure 2: *Diagram of the 3D Model extension.*

The above mechanism suggests further possibilities. For example, another new class 2DVectorModel might be added as well, and the root of an SVG ontology made equivalent to it. As it is well known, SVG is the W3C vector graphic standard whose description is available at [SVGa]. An OWL description of the underlying ontology may be found at [SVGb]. We will not explore further this option, which does not substantially differ from the X3D extension described above.

6. The advantages of merging

The newly defined 3D-CO ontology has several advantages with regard to any newly defined one:

1. It inherits from CIDOC-CRM the generality of scope: as long as it is accepted that CIDOC-CRM may deal with any cultural object, also 3D-CO can.
2. The definition is concise and easy maintainable, because changes in the parent ontologies are automatically taken into account by the URI reference mechanism.
3. The definition is easily accepted by the different communities that recognise themselves in one of the parent ontologies. To them it says: "As far as your domain is concerned, business as usual, except that there is a place where all the heterogeneous stuff is stored."
4. It allows for specialization of terms without intervening on the original ontologies. In other words, if for example a particular application needs the concept of sherd as a particular case of CIDOC-CRM `E22.Man-Made_Object`, it can be introduced in the 3D-CO without loss of compatibility, using the `rdfs:subClassOf` construct as follows:

```
<owl:Class rdf:about="#Artefact">
  <owl:equivalentClass
    rdf:resource="&cidoc;
    E22.Man-Made_Object"/>
</owl:Class>
 <owl:Class rdf:ID="Sherd">
   <rdfs:subClassOf rdf:resource =
     "#Artefact"/>
</owl:Class>
```

This trick may be used to introduce a definition that specializes a concept (Sherd as specialization of Man-Mad Object), or just make its denomination more user-friendly (Artefact).

On the other hand, this approach has disadvantages, mainly consisting in possible inconsistencies deriving from new version of the merged ("parent") ontologies. For example, imagine that a future release of CIDOC-CRM changes the denomination of E36 from Visual Item to, say, Visualization Stuff. The equivalence linkage between 3D-CO and CIDOC-CRM would hence be broken because of a formal — but not substantial — change in the parent ontology.

This is unavoidable with the OWL mechanism of mapping, and is difficult to avoid anyway. Although the embedding takes into account any update in the original, there are some critical points where change may cause disruption. This can perhaps be managed on the source side, by creating a mechanism similar to RSS, which is automatically checked and sends an alert to the children ontology administrator, subscribed to such a service, when such a potentially disruptive event takes place; or on the destination side, by checking automatically that as yet no change intervened in the version, and alerting the administrator when it happens.

Figure 3 provides a schematized representation of the process developed so far. Each cell represents a class, and different planes correspond to different ontologies. The arrow shows the correspondence as envisaged in the first option described above, that is the equivalence between X3D Scene, CIDOC-CRM `E26.Visual_Item` and 3D-CO `Model3D`.

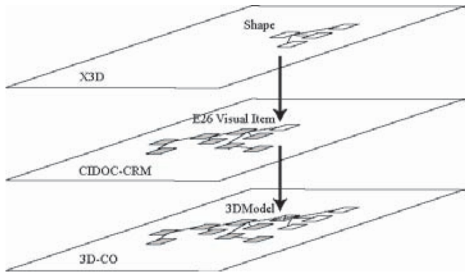


Figure 3: *Diagram of the ontology mapping.*

Figure 3 illustrates the subset of 3D-CO (represented by grey cells) that are isomorphic (one-to-one corresponding) to CIDOC-CRM.

7. Examples

7.1. Porsenna's Mausoleum

We will now show how the 3D-CO may be used to store information concerning the 3D Model of the Mausoleum of Porsenna. This is a lost Etruscan monument, of which only a description remains. This description has been used to create an X3D model, and the latter has become a testbed for different applications, see for example [Nic06], where the complete X3D description can be found.

For this example we will assume that there is a 3D-CO DTD, named CO, at the fictitious URI `www.3d-co.org`. This DTD results from the merge of the CIDOC-CRM DTD, available at the CIDOC-CRM site [CRM], and the X3D DTD, available at the X3D site [X3Db]. Thus it includes all the CIDOC-CRM entities and the X3D entities, plus any new one defined here.

For space reasons, we will not enter into further details concerning these merged DTDs, and will omit a large (and inessential) part of the description. Equivalent entities will of course be present one time only.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet type="text/xsl"
  href="3D-CO.xsl"?>
<!DOCTYPE CO SYSTEM
  "http://www.3d-co.org/DTD/3d-co.dtd">
<CO>
<CRM_Entity>3D Model 4321
 <in_class>Model3D</in_class>
 <is_identified_by> Object_ID 4321
  <in_class>
    E42_Object_Identifier
  </in_class>
```

```
 </is_identified_by>
 <has_title>
   Porsenna's Mausoleum
   <in_class>E35_Title</in_class>
 </has_title>
 <has_type>
   3D Model
   <in_class>E55_Type</in_class>
 </has_type>
...
<!- more CIDOC-CRM stuff if necessary->
<Model3D>
<!- this is the equivalent of Scene->
<!- X3D starts here ->
<Background
 groundAngle='1.309, 1.571'
 groundColor='0.1 0.1 0,
              0.4 0.25 0.2,
      0.6 0.6 0.6'
 skyAngle='1.309, 1.571'
 skyColor='0 0.2 0.7,
           0 0.5 1,
   1 1 1'/>
<ProtoDeclare name='pyramid'>
 <ProtoInterface>
  <field name='transparencyValue'
    type='SFFloat'
    value='0.0'
    accessType='inputOutput'/>
  <field name='translationValue'
    type='SFVec3f'
    value='0 100 0'
    accessType='inputOutput'/>
  <field name='scaleValue'
    type='SFVec3f'
    value='1 1 1'
    accessType='inputOutput'/>
 </ProtoInterface>
...
<!-more X3D stuff->
</Model3D>
</CRMEntity>
</CO>
```

7.2. The Bell Beaker

A similar process can be carried out with the drawing in figure 1, the bell beaker, in such a way that it is acknowledged as a Bell Beaker. Again we will assume that there is a 3D-CO DTD at the fictitious URI `www.3d-co.org`, resulting from the merge of the CIDOC-CRM DTD and the SVG DTD, available at the SVG site [SVGa].

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet type="text/xsl"
    href="3D-CO.xsl"?>
<!DOCTYPE CO SYSTEM
  "http://www.3d-co.org/DTD/3d-co.dtd">
<CO>
<CRM_Entity>Drawing 1234
 <in_class>Model2D</in_class>
 <is_identified_by> Object_ID 1234
  <in_class>
    E42_Object_Identifier
  </in_class>
 </is_identified_by>
 <has_title>
   Wiltshire Bell Beaker
   <in_class>E35_Title</in_class>
 </has_title>
 <has_type>
   B/W Drawing
   <in_class>E55_Type</in_class>
 </has_type>
 <represents>Object 9999</represents>
...
<!-more CIDOC-CRM stuff if necessary->
<!-9999 is the original Beaker->
<2DVectorModel>
<!-this is the equivalent of entity SVG->
 <has_width>523.565pt</has_width>
 <has_height>536.28pt</has_height>
 <has_viewBox>"0 0 523 536"</has_viewBox>"
  <g id="Layer_x0020_1">
    <path d="... ">
...
<!-more SVG here->
  </g>
 </2DVectorModel>
</CRMEntity>
</CO>
```

In conclusion, it is possible to create an XML integrated description of the object, keeping together both the CIDOC-CRM data — and above all the link to the original object — and the graphic data, using for the latter the most appropriate description, in our example X3D or SVG. The creation of joint DTDs is straightforward since merging operations using XML namespaces is rather common.

8. Conclusions and future work

The approach used in this paper has shown to be very promising. Its advantages have already been described, but there are many more. For instance, it is possible to store together integrated descriptions of cultural objects with 2D or 3D models, availing of internationally accepted standards, into an XML native database.

It must be noted that the exercise presented here is mainly a proof-of-concept. For example, we need to rely on resources stored in sites sometimes unavailable — to test the code it has been necessary to copy them on our server and access them from there. The CIDOC-CRM ontology refers to a previous version of the standard, and the server hosting the X3D ontology is often down. To make all this practical, provisions should be taken to guarantee safe access to resources and frequent updates — in one word, maintenance. On the other hand, the examples presented in chapter 7 show that practical applications are viable.

From the above XML encoding, domain specific descriptions may be easily extracted: writing an XSLT to extract the cultural (or graphical) information is a straightforward task.

Future work will concern the update of CIDOC-CRM OWL description, full integration of graphical standards into 3D-CO and making the resulting ontology available on-line.

Finally, an extensive number of test cases need to be fully carried out: usually problems arise from practice even when the theory looks effective and terse as in the present work.

9. Acknowledgements

The present paper has been partially supported by EPOCH, project no. IST-2002-507382. However, this paper reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained herein.

The authors acknowledge the usefulness of discussing the features of the 3D Cultural Object with colleagues of EPOCH who are working in the 3D Task Force, notably Luc Van Eicken of KU Leuven and Sven Havemann of Graz University. Although sometimes with different opinions, their expertise and their contributions to the definition of an EPOCH 3D-CO have cleared us many points otherwise difficult to overcome.

Finally, we want to thank our colleague Achille Felicetti for his invaluable technical help.

References

[Bal05] BALZER D.: An OWL representation of core definitions from the CIDOC conceptual reference Model, version 3.4.9. January 2005.

[BDN06] BEACHAM R., DENARD H., NICCOLUCCI F.: The London Charter. In *VAST2006 Proceedings — Project and Short Papers* (2006), M. Ioannides, D. Arnold, F. Niccolucci, K. Mania, (Eds.).

[CK06] CHRISTODOULAKIS S., KALOGERAKIS E.: OntologyX3D — Upper level graphics and virtual reality OWL ontology based on the X3D schema. Available at: http://astral.ced.tuc.gr/delos/content/reports/3D_Ontologies_X3D_report.pdf (accessed 18/09/06), 2006.

[Col] http://www.collada.org/ (accessed 18/09/06).

[CRM] http://cidoc.ics.forth.gr/ (accessed 18/09/06).

[Doe01] DOERR M.: A comparison of the OpenGIS Abstract Specification with the CIDOC CRM 3.2 , October 2001. Available at: http://cidoc.ics.forth.gr/docs/opengis_map.pdf (accessed 18/09/06), 2001.

[EPO] http://www.epoch-net.org/ (accessed 18/09/06).

[HPS04] HAMEED A., PREECE A., SLEEMAN D.: Ontology Reconciliation. In *Handbook on Ontologies* (Berlin, 2004), S. Staab, R. Studer, (Eds.), Springer, pp. 231–251.

[Hun02] HUNTER J.: Combining the CIDOC CRM and MPEG-7 to Describe Multimedia in Museums. *Museums on the Web* (2002). Also available at: http://www.metadata.net/harmony/MW2002_paper.pdf (accessed 18/09/06).

[ISO06] ISO 21127 - Information and documentation – A reference ontology for the interchange of cultural heritage information. Available at: http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=34424&scopelist=PROGRAMME (accessed 18/09/06), 2006.

[Nic06] NICCOLUCCI F.: Using X3D to visualize credibility parameters. In *Proceedings of the London Symposium and Expert Seminar "Making 3D Visual Research Outcomes Transparent", London, King's College, January 2006* (2006).

[OWL04] OWL Web Ontology Language Reference. W3C Recommendation 10 February 2004. Current version available at: http://www.w3.org/TR/owl-ref/ (accessed 18/09/06), 2004.

[SVGa] http://www.w3.org/Graphics/SVG/ (accessed 18/09/06).

[SVGb] http://www.pms.ifi.lmu.de/rewerse-wga1/otn/OTN2SVG.owl/ (accessed 18/09/06).

[X3Da] http:www.web3d.org/ (accessed 18/09/06).

[X3Db] http:www.web3d.org/specifications/x3d-3.0.dtd (accessed 18/09/06).

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Ontological Modelling for Archaeological Data

Andrea D'Andrea[1] Giulia Marchese[2] and Tommaso Zoppi[2]

[1] CISA- Università degli Studi di Napoli L'Orientale
[2] PIN Scarl PRATO

Abstract

The need to investigate an archaeological context causes the inevitable destruction of the upper strata in favour of the lower ones. On-site research activity therefore translates into a documentation activity (forms/photographs/surveys), which must be as neutral as possible and not necessarily biased by the scientific interests of the archaeologist. This documentation activity has recently been subjected to an in-depth analysis and evaluation. A wide spectrum of standards and reference regulations is being defined by different national and international Organisations. The various attempts made at normalizing the production of documents of excavations have lead to the elaboration of a large number of forms. The introduction of computer science in the management of archaeological records has actually complicated the picture because it has added different formats, software and operating systems, chosen by each individual researcher. Considering how dangerous it is to convert older data to newer digital formats, as is any translation from one language to another, we started a project aiming at defining an ontology able to guarantee interoperability between different archives without modifying, altering or sacrificing the archives created by each archaeologist. The primary objective of our research has been to analyse the Italian documentation produced during the stratigraphical excavation, which represents the most consistent corpus of data available also in digital format. We chose the CIDOC-CRM because it is event-oriented. In terms of content, the archaeological documentation activity may be easily schematized: it documents a past event occurring during an archaeological era and, at the same time, it documents the action of the modern-day scholar. Any excavation activity and its pertinent methodology may be easily described following this conceptual formalism.

Categories and Subject Descriptors (according to ACM CCS): I.2.4 Knowledge Representation Formalisms and Method [Artificial Intelligence]: Representation (procedural and rule-based)

1. Introduction

Excavation is an activity aimed at investigating an archaeological context, both horizontally at the open-area level, and vertically at the depth level. It causes the inevitable destruction of the upper strata (even when dealing with monumental structures) in favour of the lower ones.

Archaeological excavation is distinguished by two types of activities in the field and a successive one in the laboratory. The first two are characterized by the actions of excavation and by the documentation of all procedures carried out in the field, while the third one is performed in the laboratory, often when the excavation activity has already been completed. The documentation activity is carried out during and after the digging action, often starting with the filling of forms on field and terminating in the laboratory after the analysis of all the documents produced (photos, drawings, reports, forms and findings).

From a methodological point of view all these activities can be altered by the experience of each individual researcher, by local archaeological tradition or by the type of archaeological context investigated (tombs, monumental structures, surfaces, etc.). Thus it is possible to change the way to carry out the excavation, adjusting the methodology and research strategy to the specific archaeological area.

The archaeological survey in the field is a process involving the destruction of the data required by the archaeologist in order to reconstruct a past event. In order to avoid losing important information it is necessary to accompany all the ground investigations with other types of documentation activities. So, while an archaeologist digs, other archaeologists or technicians are involved in the "translation" of the bulk of materials coming from the excavation into written, photographic and graphic documentation.

On-site research activity therefore translates into the documentation (forms/photographs/surveys) of all the activities carried out in order to allow the archaeologists to provide all the information indispensable for analysis and study at a later time. Irrelevant information is often stored as well, as archaeologists are not always really conscious – during excavation – of the content of their investigations.

This descriptive process, called archeaography [Mob81] permits other scholars to re-use the documentation gathered during fieldwork for in-depth studies and perhaps for new hypotheses.

The documentation activity should therefore be as neutral as possible and not exclusively related to the scientific interests of the archaeologist. But as I. Hodder [Hod99] highlighted "…*whether an object in the ground has any chance of becoming an 'archaeological object' depends on the perspectives and methods of the recovery process*". All data collected on-site are necessary for the archaeologists, but only certain raw data become archaeological information through the filter of experience or scientific sensitivity of the researchers.

For these reasons the documentation activity - both for the excavation and for the documentation preceding the on-site investigation - has been in recent years the object of in-depth analysis and evaluation [RB96)].

Documentation activity of all the actions carried out in the field has a long history in the archaeological tradition and perhaps arose with the birth of modern archaeology. From ancient reports, composed of "artistic" drawings and photos, to the modern technologies used to collect digital data directly on field, the history of documentation has known many different strategies and goals. It is possible however to state that it is precisely the documentation activity to have given a scientific status to modern archaeological investigations. From the first paper reports/forms filled by archaeologists, which can represent a sort of history of the sensitivity of the archaeologists on field and their methodological strategies, we've now moved on to a sophisticated way of recording the archaeological process. I. Hodder [Hod00] has suggested using a video recording system during the excavation to improve not only the techniques of documentation of all the actions, but also to develop a "reflective" stance.

2. Standards and "Best Practices"

Nowadays, in order to contain the excavator's natural tendency to be a narrator of stories (often already present in the mind of the archaeologist), rather than a neutral excavation "technician", a wide spectrum of standards and reference regulations are being defined by different national and international organisms [ads.ahds.ac.uk/arena/links/standards.html].

The main goal of these standards is to normalize the production of excavation documents through the elaboration of a large quantity of forms. These have been designed to record the overall activities of an archaeological excavation (be they structures or objects, finds, artefacts) thus guaranteeing a further analysis of the documentation. All the forms are really accurate and designed to document in a homogeneous way the information gathered on field, without being influenced by the sensitivity and experience of the archaeologists and their scientific targets.

It is possible to distinguish various types of standards which only in certain cases may also cover the problems regarding the management of an archaeological excavation:

- **Catalogue** Standard: it determines the rules for cataloguing various objects;
- **Standards** regarding the terms used to describe an object (for example vocabularies and thesauri);
- **Metadata** Standard: it gives structure to information making it interchangeable independently from the databases used.
- **Interchange** Standard: it allows communication between computers.

The simplest and most immediate type of standard is the "thesaurus". It characterizes archaeology almost as its foundation and disciplinary premise, though we can't really speak of a general reference (and generic) vocabulary, as it is connected to specific sectors (prehistoric, classic, oriental, medieval archaeology, etc.) which are divided into sub-disciplines that further specialize the different thesauri (Etruscology, Greek archaeology, the Iron Age, etc.). At this level the primary goal of a standard is to guarantee "homogeneity" in the description of any object, a sort of single description protocol with values defined by pre-fixed lists (vocabularies, thesauri, dictionaries) chosen by each archaeological team according to specific needs.

Many standards devised especially for the management of museum collections exist [http://www.fish-forum.info]. These are often implemented to guarantee a rapid inventory of the objects found during an archaeological investigation in the field. Thanks to widespread computerization, these standards allow museums to be efficiently automated and kept up-to-date for any future developments in the sector.

Notwithstanding the release of a number of standards, often widely distributed at the national level, the archaeologist's work in the field has not changed substantially. On the contrary, the university teams, often released from the obligation to produce and deliver "standard" documentation, make use of a variety of different methods and systems for the recording of information.

For this reason, the initial goal of the documentation activity, that is to guarantee a sort of scientific interoperability of data, failed. Nowadays there are too many different standards and information systems that compete against each other, all aimed at recording the archaeological excavation. All the attempts made to "create" an interoperability about the structures and the content collided with many, problems not easy to solve: it is impossible to figure out a sort of super-form able to synthesize all the various forms. In the same way it is hard to foresee a single dictionary/thesaurus/vocabulary on which all the archaeologists agree, as it is to oblige archaeologists to use a specific system or format or platform. The design of a unique methodology effective for all the researchers should also be foreseen.

The recent introduction of computer science has complicated the picture because it added different formats, software and operating systems, chosen by each individual researcher.

Although the passage from Relational Databases to "mark-up language" made the data normalization process more flexible, thus more "aware" of the quality of the information, rather than the quantity, it hasn't always been possible to understand how the researcher's ability has lead him/her to the classification/interpretation, thus to the formalization of the data. This difficulty emerges even more clearly if one thinks of the need, common to most archaeological investigations, to re-use previously compiled documentation, often formalized according to systems and excavation methodologies which differ from those based on the principles of archaeological stratigraphy. Each new archaeological hypothesis will be built upon the documentation previously acquired during on-site investigation. It is easy to figure out how each historical and graphical (3D or virtual) reconstruction is influenced by the quality of the recorded documentation.

For these reasons, in order to avoid the implementation of a new system incapable of communicating with other systems and archaeological archives, we decided to experiment the use of the knowledge-based system in order to understand and formalize the conceptual model underlying the design of the archaeological documentation.

The absence of interoperability, caused by the diffusion of different platforms, programs and formats, is determined mainly by a profound semantic diversity, only partially solved by the definition and circulation of thesauri and thematic dictionaries. Perhaps what still impedes a real and true interoperability between on-line resources, the portals and thus the data is the absence of a clear formal representation of knowledge-based models that are at the basis of the normalization and computerization of information.

2.1. Italian ICCD Standard

The reference model for cataloguing the Italian archaeological heritage was published in 1984 by the Istituto Centrale per il Catalogo e la Documentazione (www.iccd.beniculturali.it), an institute of the Ministry of Cultural Heritage. It refers to a limited number of forms: Stratigraphic Trench (SAS), Stratigraphic Unit (US), Archaeological Find (RA), "Wall Covering" Stratigraphic Unit (USR), Paleo-Anthropological Remains and Archaeological Monuments (MA).

As one may read in the foreword of the Normative that ICCD released in March 2004 (*Sistema Informativo Generale del Catalogo* - Normativa 3.00) [Nor04]: "*…the intense job of systematization of the entire catalographic process in its methodological and operative aspects… has entailed an accurate revision of the most frequently used catalographic forms and of the regulations regarding their compilation… To these requirements, of a technical, practical and operative nature, linked to the refinement and natural evolution of a normative occurring in its emerging aspects over a ten-year period of time, one must add those made necessary by the current and delicate phase of data diffusion which must be reconciled with the principles of privacy, safeguard and intellectual property; such demands have rendered necessary the insertion of appropriate fields in which the institutions responsible for the catalographic procedure are necessarily involved regarding the "sensitivity" of the information and the consequent differentiated access according to the user profile*".

The activity of the ICCD however is not limited to the definition of forms, but also of the format standards in order to guarantee an optimal management of all the available resources. Through the years, along with a consistent investment of resources for the planning and implementation of forms to be compiled according to formalized ways of describing objects (vocabularies), a reflection regarding format standards has also developed. In reference model 3.00 [Nor04] there is a section reserved for multimedia applications standards, which underlines the advantages of using widely available commercial products and which are characterized by the use of interchangeable formats (*.tiff, *.jpeg, *.dwg, etc.) that allow for some interoperability between different sources.

Notwithstanding the coordinating action of the ICCD in promoting and favouring the adoption of format standards (dictionaries and forms) and of digital support (file types), the activity of documentation in universities has moved along different lines due to widespread computerization of the archaeological excavation. Proprietary programs have been implemented for specific research projects (GIS solutions, Database, Multimedia tools, etc.). The absence of a common framework determines an important issue not easy to solve: the migration and the integration of these proprietary data and "standards" in a unique system. This situation becomes more complex (and difficult to avoid and overcome) when different teams, using a variety of recording systems and standards, work in the same archaeological areas. Unfortunately in this case, as in other similar conditions, the integration is carried out at the level of historical synthesis and not of raw data. So it is impossible re-use this *oriented-data* in order to propose new historical reconstructions and hypotheses.

2.2. Syslat

Alongside standard systems and national regulations for the formalization of archaeological documentation, a number of informatics systems are available, often implemented by a single excavation or research team. Their objective is to simplify the process of recording, filing and researching of the information stored in the document archive (forms, images, drawings, etc.).

During the excavation of Cumae, the most ancient western Greek colony, the University of Naples l'Orientale adopted the Syslat registration system since 1994.

Originally conceived as a tool to record excavation data of the proto-historic site of Lattes (Montpellier – France), since 1984, when it was first experimented, Syslat (Système Lattes) underwent several releases that progressively assimilated the suggestions deriving from use. Outside France, the system has been used in the investigation carried out by the University of Naples l'Orientale in the ancient Greek site of Cumae. The experimentation created a portable system provided with a wide set of personalization functions, which can be configured according to the needs of any excavation.

Syslat is not only a translation in digital format of the stratigraphical excavation model formalized by Harris. Even if it is based on the fundamental principles of stratigraphy, it is a tool for a "guided" organization of field data (from recording of actions to graphical and photographic documentation and sample collection) and for an integrated management of all data, including statistic-quantitative analysis of materials.

Syslat represents a great container in which appropriate scripts allow access to recorded information: from stratigraphical units to "fact" and "set" records, from photos to graphic archives, from quantification records of materials to the typology of individual ceramic finds and so on. The global archive is structured in five different modules (terrain, objects, samples, documentation, utility) from which additional hierarchical sub-levels may be accessed.

The system includes an ample iconographic dictionary to classify ceramics and several vocabularies guide the operator in compiling forms. Syslat also has a module to personalize the database and to add new definitions to glossaries and dictionaries.

Some years ago we realized the porting of the tool into an interoperable framework, converting all data gathered during the excavation of Cumae in XML, and we implemented a customized management system [DN02] using an open-source program.

3. A Preliminary Ontology for Archaeological Data

3.1. Introduction

By comparing the two standards (ICCD and Syslat) we immediately found out that, apart from a few items and a kind of in-depth fields for the description of the archaeological context, these two systems were different for the conceptualisation and organisation of data: for instance, the numeration system assigned to the stratigraphical units. While the former is similar to the traditional Harris system, with a numeration not linked to the interpretation of the excavation, the latter vice-versa relies on the sub-division of the investigated area into "zones" corresponding to the ancient organization of the site (blocks, quarters, streets, functional areas).

Moreover these two systems used different vocabularies (one in Italian and one in French) and many different items corresponding to the sensitivity and the experiences of each institution. For example while the ICCD system offers a greater wealth of descriptions for the identification of stratigraphic relations, in particular of wall structures. Syslat is based only on anteriority, posteriority and contemporaneity, the three fundamental characteristics of the relations between US.

A further element of differentiation is given by the aggregation of US into groups. The ICCD is based on the grouping of archaeological levels according to functional units, which should correspond to sets of actions and activities referable to a single function (habitation, house, private/public building, sacred area, etc.). Syslat introduces instead an intermediate element, the fact, whose function is to schematize the construction process of the matrix based on the simple observation that some actions, although split amongst more than one US, represent in any case a single action: for example, trench and fill, wall and foundation, etc.

Considering how varied the methodology used to investigate an archaeological area is, we immediately abandoned the idea of integrating the metadata of the two systems into a single database. The differences were in the conceptual models used to dig and to collect data, which followed a different strategy, the former technical, the latter interpretative, and not in the software or in the design of database or the management system.

Furthermore considering how dangerous it is to "convert" older data to newer digital formats, as is any translation from one language to another, we started a project aimed at defining standards that guarantee interoperability between different archives without modifying, "altering" or sacrificing the archives created by each archaeologist. We wanted at the same time to maintain not only the data and the documentation collected, but also the methodology utilized in order to better understand how the data was created and then processed.

Consequently our project does not consist of recognizing and underlining the fields in common with the different structures present in the forms (for example: place, location, period, chronology, phase, stratigraphical relations, etc.) for it might appear as a simple definition of metadata. Instead it consists of "extracting" from each definition a representation of the conceptual model the archaeologist has referred to during his/her fieldwork and which translated into documentation. The objective is not limited to attempting the integration of multi-temporal and stratified databases, but rather to the necessity of "comprehending", thus of "representing" in a "transparent" manner, the processes carried out by the archaeologist in his/her own knowledge-based domain (stratigraphical excavation, open-area excavation, in-depth trench excavation, artificial strata excavation).

To guarantee interoperability between different repositories (formalized using different standards) the first step should be to have a common standard. It is impossible however to force archaeologists to use the same protocol or standard, thus sacrificing their point of view.

To avoid constructing a new standard for the entire community of archaeologists (a super-standard), we decided to work on a new approach using an ontology to understand the work of the archaeologist on field. Other proposals, with our same scope, failed because they were based on the idea of a common distributed infrastructure to guarantee the interoperability among archives [http://www.progettodice.it].

The ontology was deemed particularly useful to formally "describe" the archaeologist's activity; the ontology is synthetically defined as formal cognitive models in a certain domain. In this perspective it is possible to implement a system based on the common ontology in order to integrate different repositories created using Database or other systems into one management system.

We decided to formalize these methodologies using the CIDOC-CRM in order to highlight the semantic connections between these two recording and documenting systems.

3.2 CIDOC-CRM

The CIDOC-CRM (http://cidoc.ics.forth.gr) [CDG*05] is an ontology created in order to offer "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 used 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 those published by museums, libraries and archives". After about 10 years the CIDOC-CRM has become an ISO recommendation, while a definitive elaboration is expected from the technical committee ISO/CD 21127.

Even though the CIDOC-CRM is composed of over 80 classes and 110 properties, it may reduce to a "light" meta-scheme (diagram) that has at its core a correlated event, through bi-directional property, to other classes: Object, Subject, Place and Time. Such a simplified scheme (diagram) works with a rapid integration of data with different formal structures. So CIDOC-CRM is a sort of *Top-level ontology* describing very general concepts like space, time, matter, object, event, action, etc., which are independent of a particular problem or domain [Gua98]; different kinds of ontology according to their level of generality.

Although this knowledge-based model has been implemented especially for the documentation of museum collections, CIDOC-CRM may also be adopted to describe the documentation gathered in the course of field investigations. We chose the CIDOC-CRM because it is *event-oriented.*

In terms of content, the archaeological documentation activity may easily be schematized: it documents a past event occurring during an archaeological era (archaeological strata) and, at the same time, it documents the action of the modern-day scholar. Any excavation activity and its pertinent methodology may be easily described following this conceptual formalism.

3.3 Mapping of Archaeological ICCD and SYSLAT Forms into CIDOC-CRM

The main peculiarity of our project was based on *a highly interdisciplinary approach* playing a fundamental role in analyzing the structures of the different recording system.

Even if the two systems were based on the same vocabulary, there is no guarantee that they can agree on certain bits of information unless they commit to the same conceptualization. As each recording system has its own conceptualization, a condition that is necessary in order to make an agreement possible is that the intended models of the original conceptualizations overlap.

Each archaeological method or strategy can simply be described as a task-ontology which, as object of its activity, produces a specific documentation. *A task-ontology* describes generic task or activity by specializing the terms introduced in the top-level ontology. In our case it was possible to illustrate the two different systems by considering the selected strategies as a specialization of the concept of the Stratigraphical Excavation.

So it is impossible to verify the documentation without knowledge of the method adopted by the archaeologist. Similarly it is impossible to guarantee the integration of different data acquired in the same area by different archaeological teams.

The primary objective of our research has been to analyse the documentation gathered during the stratigraphical excavation, which currently represents the most consistent corpus of data available also in digital format according to national regulations. An experimental project regarding the mapping of excavation forms according to the CIDOC-CRM standard has been initiated and its primary role is "*to enable information exchange and integration between heterogeneous sources of cultural heritage information*".

We've used CIDOC-CRM in order to supply the data with the semantic definition required in order to transform the archaeological documentation into a coherent global framework.

Our project is based essentially upon recognizing the correspondence of meaning and content between different classes of CIDOC-CRM and the fields defined in the forms chosen as samples for this experiment. In synthesis our objective is to propose an extension of the CIDOC-CRM classes and to pinpoint as many sub-classes and/or relationships for the structure of certain excavation forms (which are standard on ministerial forms and on those forms created by a university team).

As defined by CIDOC-CRM, extension signifies that CRM classes subsume all classes of the extension, and all properties of the extension are either subsumed by CRM properties, or are part of a path for which a CRM property is a shortcut. Thus the model may be enriched and "customized" without losing "compatibility" with CIDOC-CRM.

The model we chose does not limit itself to a simple description of the archaeological "categories", but rather to a detailed one, since it inherits all the sub-classes and relationships of CIDOC allowing a more sophisticated formal analysis of the recorded data.

This research – linked to a larger project named AMA carried out in the framework of the European Network EPOCH (www.epoch-net.org) – relies on a previous research having similar goals, among others the one concerning excavation forms from the archaeological site of Cumae [CDF02] [DN02].

The Cumae forms, converted in XML, have been edited in RDF and visualized with the Protégé, an extensible, platform-independent environment for creating and editing ontologies and knowledge bases [protege.stanford.edu].

The graph shows a plane structure without relations and with circular-type references. According to this method of structuring data, all information is placed on the same level, thus no hierarchy is visible.

In order to extract the relationships between the data, we have extended the ICCD and Syslat forms beginning from an initial analysis which has lead to the "conversion" of certain types of information into classes and of others into properties.

Analyzing the ICCD forms, it appears that several fields, such the Cumae forms, are placed according to implicit relations not included in the form structure. The most evident "anomalies" for instance concern chronology/dating and author. Dating may refer to the archaeological event, to the excavation or to the documentation. In a similar way, the author may be the person who produced the event in the past, the excavator or the form compiler. In fact the form label simplifies its understanding, but a simple conversion shows, as in the previous case of Cumae, how difficult it is to extract a semantic structure from the fields. A difficulty clearly emerges – showing a definite impact on the quality of documentation – regarding the field "stratigraphical reliability": does it depend on the nature of the excavation (for example an investigation performed with a mechanic device provides less data than a brush-made one) or on the archaeologist's perplexity about the interpretation of the stratigraphy (basing on experience/competence)? Neither appropriate metadata system is sufficient. Why did the archaeologist chose one system (the mechanic device) instead of another (the brush)? Possibly he/she was interested in reaching quickly the lower layers and considered as irrelevant the sacrifice of a layer in the overall economy of the excavation management and interpretation. Often such information is located in the "interpretation" field, a sort of "black hole" where the archaeologist puts the most important data for the laboratory reconstruction of the excavation. In a similar way, the attached documentation (photos, drawings) is compressed by the methodology chosen.

Another aspect emerging from the semantic reading of the US form derives from the so-called stratigraphical relations, that is the spatial and temporal relations among different Stratigraphical Units (natural/anthropic actions discovered during the excavations and individually numbered). In this case they are not class attributes, but spatial relations among US forms which sometimes document events that happened before/after the one recorded in the form.

This preliminary analysis produced a first manual mapping consisting in the recognition of the correspondence between the classes of the US form and the entities of CIDOC-CRM. In this case, using Protégé we created a sub-class corresponding to the US form fields for every single entity.

The following tables show a synthetic exemplification of the draft predisposed for highlighting the correspondences between our forms and the CIDOC-CRM entities.

ENTITY-TAG	**CONCEPT**	**CRM-ENTITY**
SCHEDA_US	US positive form	*E31_document*
NOTES	Each unit in the excavation has a corresponding US form. The tag refers to the database report or, alternately, to the paper sheet concerning the unit.	
RELATION	**WITH**	**NOTES**
P70_documents	AIE_13_us	
P4_has_time_span	AIE_06_anno	
P70_documents	E5_event	Event: excavation

ENTITY	**CONCEPT**	**CRM-ENTITY**
DIRETTORE	Director of excavation	*E82_actor_ appellation*
NOTES	Name of the director of the excavation. Tag used in forms: US, USM, USN.	
RELATION	**WITH**	**NOTES**
P11_participated_in	E5_event	E5_event (excavation) P108_has produced:
		scheda_us
P49_is_former_or_ current_keeper_of	scheda_us scheda_usm	

PROPERTY	**CONCEPT**	**CRM-PROPERTY**
ANTERIORE_A	Before to (US, USM, USN)	*P120_occurs_ before*
NOTES	Chronological relationship with positive and/or negative stratigraphical unit. It refers to the USM, US, USN chronologically and immediately before to the unit in exam. Tag used in US, USM, USN forms.	

Using the Protegè editor and an RDF file present on the CIDOC-CRM website, we subsequently created sub-classes corresponding to the ICCD and SYSLAT standards for every CRM class singled out. It was thus possible to maintain the sub-classes and properties defined by CIDOC-CRM.

After checking the correspondences between our items and the entities of CIDOC-CRM, the mapping relied on the definition of possible new sub-classes and sub-properties. It often isn't sufficient to find the correspondences to highlight the semantic content without pointing out the paths or links that join each field/entity to one another. In Figure 1 we show how it was possible to link the compilation of the form to the author avoiding to link this task to the author of the excavation.

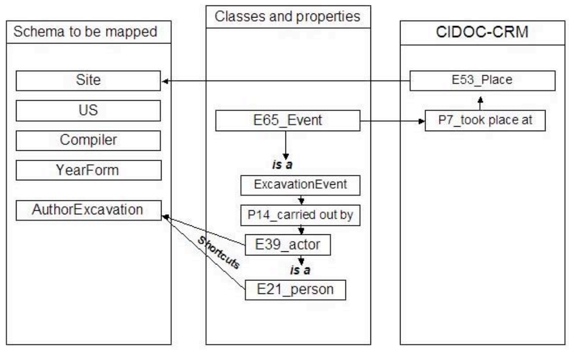


Figure 1: *Mapping mechanism.*

At the end of the process of analysis, we created the mapping procedure and generated the RDF file with the protege editor. Figures 2 and 3 show the hierarchy and the relationship among the classes CIDOC-CRM and those of the Italian standards.

The mapping thus determined is now available in the information management system of the excavation of Cumae, created as a re-adaptation of the old Syslat system in the open-source Exist environment [Fel06]. The data may now be researched and visualized in the proprietary format or exported in RDF in a format that complies with CIDOC-CRM.

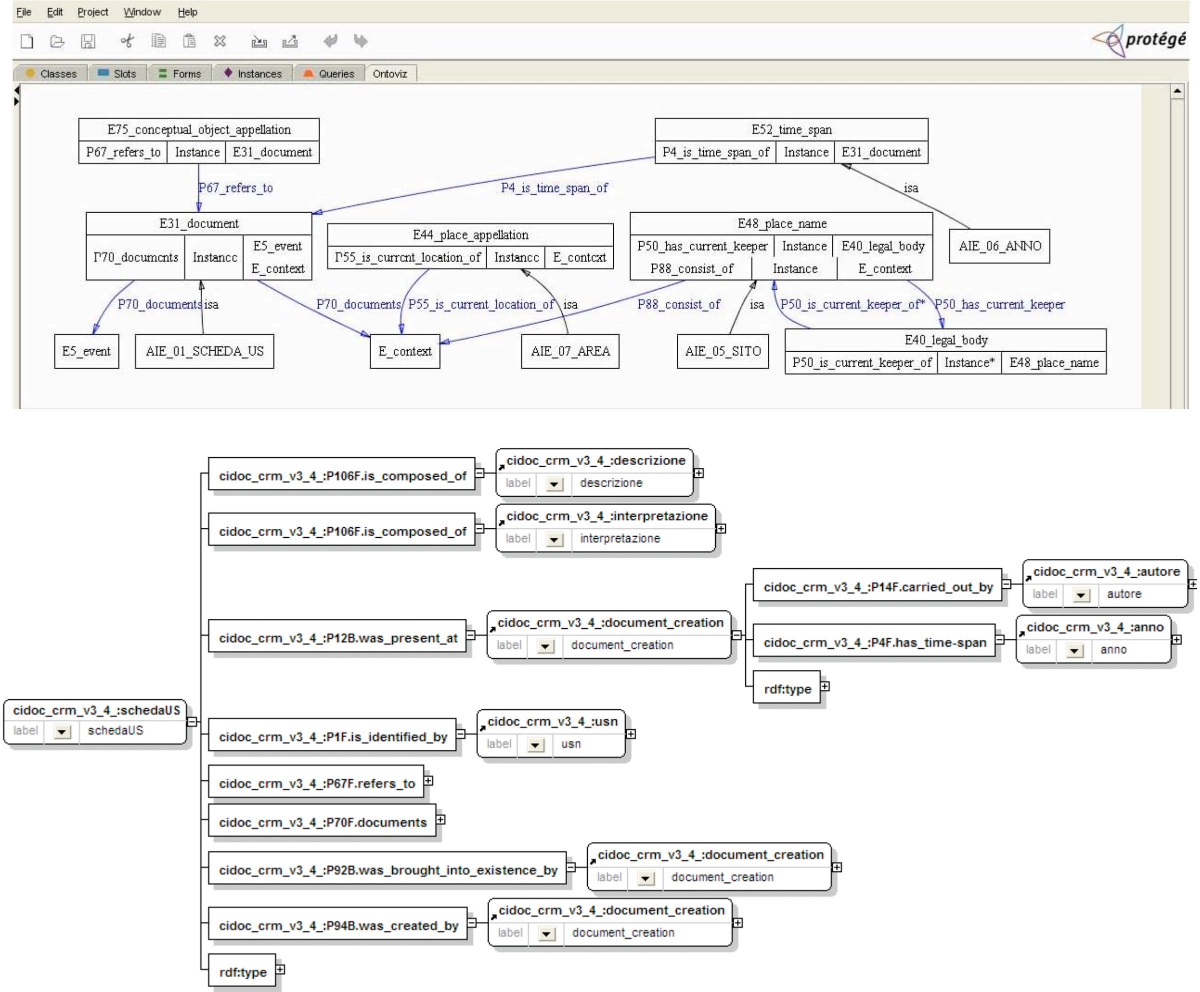


Figure 2 and 3: *The Graphs represent the hierarchy between the new sub-classes and CIDOC-CRM classes.*

4. Conclusion

The use of standards in the description of archaeological activities, both at the level of excavation and of the cataloguing of goods and handcrafted artefacts, represents a strong demand that cannot be limited, as in the case of the Italian ICCD, to ensuring a correct conservation and management of the archives, especially in the field of the management of archaeological goods. To this day a common descriptive system and a standard international vocabulary are lacking. There are difficulties turning the various attempts made into a single protocol.

The existence of national and/or local regulations and systems further complicates a picture already made complex by the habit of archaeologists to "create" new description criteria and new formalizations. As we have seen, the introduction of computer science hasn't simplified the situation. It has instead made it more articulate because of the unlimited possibilities to create and adapt local and temporary systems, which do not have the pretence of becoming object of new studies for new hypotheses.

The problem of standards, from whose solution a new, more scientific relation with the archaeological documentation - which represents the only testimony of the activity on the field - may result, does not appear to be an immediate necessity.

The benefit of using CIDOC-CRM is evident: on the one hand it forces archaeologists to conceptualize and formalize the excavation methodology and the recording of the data collected in the course of field work; on the other hand it allows the same researchers to maintain their systems guaranteeing the possibility for other archaeologists to check and re-use the data through CIDOC-CRM mapping.

Standard in our mind signifies primarily having the possibility of verifying the quality of the data in a discipline that, contrary to hard-sciences, does not allow the repetition of such a destructive experiment as the excavation. Furthermore it is possible to control the procedure and, no less important, to provide the framework for a better planning and implementation of the protocols chosen at the local level.

The disadvantages (minimal, if compared with the advantages) are to define in a conceptual and formalized way, according to a modality that is often uncongenial to the archaeologist, the methodological and theoretical itinerary that guides the researcher in his/her way in their process of cultural interpretation of the facts, starting from the material and historically reconstructed culture.

From a methodological point of view, the work carried out so far seems to align itself with what has already been experimentally accomplished by the Centre for Archaeology of English Heritage [CGF*04] which, as declared in the foreword, is based upon the attempt to model a conceptual frame for all the archaeological data created by man. Rather than by an extension, the mapping in this case has been carried out by simplifying the CIDOC-CRM classes.

In the future our objective will be to formally "describe" other types of forms and/or document representation in order to test the potential of such an approach and to guarantee the interoperability between information pertaining to an archaeological context investigated over a long period of time.

In conclusion our scope was to understand the conceptual model used by archaeologists to record data excavation in order to propose a new form for the documentation according the semantic structure supplied by CIDOC-CRM. Only this way will it be possible to obtain real data integration and a re-use of data without altering and sacrificing the specific background and targets of each individual archaeologist.

The experimentation carried out during our project may represent an option for the creation of an ontology for the cultural heritage. Using CIDOC–CRM has definitely caused problems of adaptation and conversion of our requirements, specifically those regarding research, rather than those regarding museum conservation guaranteed principally by CIDOC-CRM. The risk would have been to formalize a new ontology for cultural heritage freed from the CIDOC-CRM. Its achievement would have certainly been simple without having to go through the "choices/decisions" container in the CIDOC-CRM not easily extendable to the archaeological area. However we would have created another new standard, further complicating a picture that today seems to be characterized by incommunicability.

Facilitating the exchange of information, on the contrary, allows an improvement of knowledge, the safeguard of local and world archaeological heritage.

Acknowledgments

Work on the global mapping system is presently carried on within the project AMA, by a partnership formed by CISA, CIMEC, IAA, Oxford ArchDigital, PIN, ROB, University of Kent, University of Oslo, UNIREL, and VARTEC.
AMA is developed in the framework of the EU funded Network of Excellence EPOCH (http://www.epoch-net.org/). EPOCH is funded by the European Commission under the Community's Sixth Framework Programme, contract no. 507382. However, this presentation reflects only the authors' views and neither EPOCH, nor the European Community are liable for any use that may be made of the information contained herein.

References

[CDF02] CRESCIOLI M., D'ANDREA A., NICCOLUCCI F.: XML Encoding of Archaeological Unstructured Data. G. BURENHAULT (Ed.), *Archaeological Informatics: Pushing the envelope*. In *Proc.* of the 29th CAA Conference, Gotland April 2001, BAR International Series 1016, Oxford, pp. 267-275.

[CGF*04] CRIPPS P., GREENHALGH A., FELLOWS A., MAY K., ROBINSON D.: *Ontological Modelling of the work of the Centre for Archaeology*, September http://cidoc.ics.forth/docs/Ontological_Modelling_Project_Report_%20Sep2004.pdf)

[CDG*05] CROFTS N., DOERR M., GILL T., STEAD S., STIFF M.: *CIDOC CRM Special Interest Group*, Version 4.2, June 2005 (http://cidoc.ics.forth.gr/docs/cidoc_crm_version _4.2.pdf)

[DN02] D'ANDREA A., NICCOLUCCI F.: Database di scavo ed Internet. L'accesso remoto all'archivio Syslat. B. D'AGOSTINO, A. D'ANDREA (Eds.), *CUMA. Nuove Forme di Intervento per lo Studio del Sito Antico.* In *Atti della Giornata di Studio, Napoli 12 Febbraio 2001*, AION ArchStAnt, Quad. 14, Napoli 2002, pp. 167-175.

[Fel06] FELICETTI A.: MAD: Managing Archaeological Data. Paper presented at *7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2006), Cyprus.*

[Gua98] GUARINO N.: Formal Ontology and Information Systems Formal Ontology. In *Information Systems, Proceedings of FOIS'98*, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15.

[Hod99] HODDER I.: *The Archaeological Process*. Blackwell, Oxford.

[Hod00] HODDER I. (Ed.): *Towards reflexive method in archaeology. The example at Çatalhöyük,* MacDonald Institute for Archaeological Research, Cambridge.

[Mob81] MOBERG C.A.: *Introduzione all'Archeologia*. Milano.

[Nor04] Normativa 300 (2004) http://www.iccd.beniculturali.it/standard/normative_300.html

[RB96] RENFREW C., BAHN, P.: *Archaeology: Theories, Methods, and Practice*. 2nd Edition, Thames and Hudson, London.

All the references have been visited on 10-07-2006

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Evaluating the Social Context of ICT Applications in Museum Exhibitions

L. Pujol Tost[1] and M. Economou[1]

[1]Department of Cultural Technology and Communication, University of the Aegean, Mytilini, Greece

Abstract

The spreading of Information and Communication Technologies (ICT) in exhibitions is, among other reasons, due to the wish of curators to find new ways to improve visitors' experiences in museums. This has lead to an interest to understand if and how they really work as a museological and museographical element and assess their effectiveness. However, systematic studies in this field remain very limited or are not of sufficient depth. Although the technological field has a long tradition of assessment, this usually concentrates on technological, attitudinal or cognitive issues and does not take into account the specific features of the visit in a museum or cultural heritage site and the importance of the social context. This paper stresses the need to carry out and take into account the results of a systematic body of analyses dealing with how technological displays are really used. It also discusses the need to concentrate on the social dimension of the visit and use of ICT and to develop the methodological aspects. Based on previous studies and on our own research, the contribution of this paper is twofold: firstly, it provides an overview of empirical results concerning the use of different kinds of ICT exhibits and secondly, it discusses some preliminary ideas aimed at the construction of a methodology for evaluation. The aim is to establish the basic guidelines for the effective integration of ICT applications in museums and cultural heritage.

Categories and Subject Descriptors (according to ACM CCS): J.5 [Arts and Humanities]: Cultural Heritage, Museums, evaluation, qualitative studies

1. Introduction

The spreading of new technologies as communication tools in exhibitions is not only due to fashion and socio- economic pressures, but corresponds also to the most recent stage of a museological renovation trend which started in the last third of the 20th century and aims at improving the visitors' experience. As ICT applications have proven to be effective for learning and communication in other contexts, museums adopted them for their exhibitions hoping to introduce better ways of communicating with their visitors and to encourage their participation. This interest in their integration in exhibitions has now led to an interest in verifying if they really are effective in the museum environment and understanding how they operate in that context.

Three fields have undertaken research evaluating ICT applications in cultural heritage. The first of them is technology/engineering: this has a long tradition of studies, for example inside the specific branch of Human-Computer Interaction, but mainly focused on technological issues related to usability of the interface. As the field evolves, research studies have recently been concerned with some aspects closer to museum interests, like multi-user environments and interfaces, but they still do not question the traditional linear and sequential computer interaction paradigm. The second field is formal learning environment, which is concerned with cognitive issues arising from the use of technology for learning, but their results can only be extrapolated to some extent to museum settings because the contextual conditions of the classroom and the exhibition are different. The last field where evaluation studies have taken place is museums, where the spreading of ICT displays has recently led to an interest in evaluating their effectiveness from a communicative/learning point of view. In this direction, cultural organizations have often collaborated with external bodies or commissioned related studies. However, these studies are either mainly aimed at studying visitors' attitudes and perceptions [e.g. VKT*01, OBP05] or they are undertaken by organizations and researchers working in the previous two fields and for this reason they are more concerned with their interests, rather than taking into account museological issues, and especially the social dimension of the visit, which has been demonstrated to be paramount in the informal learning environment.

According to recent museological thinking, the museum experience can be conceptualized as the dialectical relationship each person establishes with space and time [MW05], that is, with all the components constituting this physical, emotional and cognitive environment. ICT applications seem to fit perfectly this experiential perspective because they necessarily entail an interaction, a dialogue, between the machine and the user. The problem is that until recently this dialectical concept has been understood as an individual experience -an idea shared and reinforced by the technological interaction paradigm- while, the education and later the museum field [McM88, FD00, HG94, FD92] have shifted their attention to the social experience, taking into account how people interact with each other. In situ observations have shown that even in the case of individual visits [Gal03, vLHH05, vLHH02, HvLO05], people always explore exhibits through direct or indirect interaction with all the visitors that are sharing the same space: the importance of this "social dimension" as a crucial element of the museum experience has been demonstrated by the fact that even diametrically opposite approaches to the experience of art in museums reach the same results when the construction of meaning is achieved through social interaction [Pie05].

Thus, there is a need to carry out evaluation studies which take into account the specific features of both exhibition visit and technological applications, and develop the data collection and analytical methodology accordingly, because as some authors have pointed out [HvLO05, Rou04, Man01], the traditional quantitative methods (e.g. timings, learning tests) have proven to be insufficient for offering a real understanding of what really happens with high-tech exhibits in museum galleries and study the complex set of parameters which affect the visitors' experience.

2. Building an integrated methodology for ICT evaluation in museums

Evaluation should be developed in two complementary directions: from an empirical point of view, focusing more on the way people behave at and around ICT applications, including the qualitative aspects, and not only on usability or cognitive results; and from a methodological point of view, trying to identify standard indicators of the analyzed phenomenon and then look for the factors which determined its existence. This might seem to some extent contradictory -as typified indicators are associated with the behavioural model of evaluation, examining mainly external, measurable indicators of learning- but we have to bear in mind that observable does not necessarily have to be associated with quantitative methods and learning. In order to understand the impact of technological exhibits in any of its perspectives (knowledge acquisition, emotional aspects, social interaction, etc.) we need elements of analysis which have to be necessarily external (when the researcher's point of view is used) or externalized (when the visitor's point of view is used).

Starting from this fact, different evaluation methods can be used (observation, questionnaires, interviews, focus groups, etc), each of which can be aimed either at qualitative or quantitative factors because these features are not exclusively inherent to any specific evaluation category but depend on the goals established by the underlying museological/learning theory. In our research we are concerned with the development of an observational methodology because it is one of the most powerful means to describe or verify fundamental hypothesis about what happens during interaction with exhibits -namely, the process of constructing meaning through social interaction in the social-constructivist model of learning. In conclusion, we propose the development of an integrative methodology because, as other authors have pointed out, the dichotomy between quantitative and qualitative approaches is erroneous [HG94, Hei82, Mac93], while in most cases a combination of methodologies might be more effective to obtain the whole picture of the element under analysis [vLHH02].

Some recent evaluation studies of technological displays have used observation in order to study different aspects of visitors' engagement and learning. They are relevant because not only they deal with qualitative evaluation but they are also explicitly aimed at building a specific methodology. The first one was intended to design and implement a 3D projection about Astronomy for a Science museum in order to analyze its relationship with engagement [Pod04]. To that end, the researcher conducted observations of 14 sessions and interviewed 10 of these groups. The first observations helped to identify a range of behaviours indicating engagement or the lack of it and were classified along two axes: physical/verbal and active/passive. Afterwards, he made hypotheses about which factors could influence engagement and verified them through interviews.

The second study tried to verify if technology allowed group interaction and learning in a temporary mixed reality exhibition organized at the Nottingham Castle in the UK [NG02]. The Storytent was composed of two screens forming a tent and allowing the projected images to be seen both by the visitors who controlled the navigation from inside and by those who were outside. Again, the researcher conducted direct observation in order to identify and verify if the learning indicators foreseen by socio-constructivist theory were present. She observed different activities belonging to collaborative exploration of exhibits, such as turn-taking for interaction, pointing at objects; verbal communication, and the adoption of a leader's role when interacting with the exhibit while the rest of the group observed. Starting from that, four kinds of behaviour were observed, which indicated indirectly that a learning process was taking place: storytelling, often from mothers to their children; asking questions about the castle; relating the information presented at the exhibition with previous knowledge of historical facts; and relating digital artefacts with real objects.

The relationship between interaction and learning has also been analyzed in experimental situations with virtual environments by authors who follow a constructivist model, supporting the idea that learning is a process of meaning construction which is carried out through interaction with the environment, the contents and the partners [Rou04]. Accordingly, they put the emphasis on the development of the task rather than on its results. Through observations, interviews, questionnaires and task resolution tests, it has been deduced that learning can be tracked, from an individual point of view, through conceptual change, additional knowledge and changes in behaviour; and from the social perspective, through verbal interaction (asking questions, explaining the contents, connecting with previous knowledge or the surroundings), collective decision making, conflict resolution and peer teaching.

We applied some of the lessons learned in these studies to the evaluation carried out at an exhibition organised at the Trajan Markets of Rome by the Istituto per le Tecnologie Applicate ai Beni Culturali (ITABC) from September 15 to November 20, 2005. The exhibition was called "Immaginare Roma Antica" and presented a selection of the different applications related with the ancient city of Rome, the Roman Empire or innovative research implementations submitted to an international call for technological applications (VR, MM, audiovisuals, etc.) inside the Virtual Heritage Centre project promoted by the Rome City Council, the Imperial Roman Forum Museum, UNESCO, Region of Lazion Funding Group, the Italian National Research Centre and LUISS University. This exhibition, in which audiences were able to interact with different high-tech exhibits, offered the invaluable opportunity to undertake a survey aimed at assessing the visitors' perception about the use of ICT in the Cultural Heritage field and the way different kinds of technological displays are used in the informal learning context. The project is presented in more detail, together with the preliminary results of the evaluation, in the Projects and Short papers volume [FPP06]. The social and qualitative aims of the evaluation were reinforced by the fact that the content of the ICT applications were also different and therefore the comparison of the effectiveness of different interfaces could not be as conclusive as it would have been if it had been in an experimentally controlled situation in which the same content would be tested using different software and hardware.

The evaluation included interviews and observations, which we undertook in order to gather qualitative information about real situations and compare it with the visitors' answers in order to clarify or explain them. We carried out two different kinds of observations: staying for forty minutes at each exhibit, and tracking one example of each visitor category along the visit. In both cases, we wrote down in a standardized sheet all comments and behaviours in relation to four different groups of variables: individual, social, technological interaction and learning. In relation to the first three categories, we distinguished between social exchange and interaction with the exhibits because, as we mentioned above, this research is based on the premise that the experience of the visit comprises of the mutual influence of three different elements: the exhibits, the visitors' personality and previous experience and their behaviour. The indicators related to the visitors alone were divided into social and individual behaviours and organized following a gradation. Individual behaviour included: just take a look, read/look at it, observe what others do, talk to other visitors, and individual interaction with technology. Social behaviour included: just take a look, read/look at it, observe what others out of the group do, interaction within the group, talk to other visitors, individual interaction with technology while the rest just look, turn-taking, collaborative interaction with the technology (one uses and the rest help), collaborative interaction with the technology (all use at the same time).

These indicators have been used in the substantial body of research concerned with visitors' behaviour analysis which follows the ethnomethodological approach [vLHH05, HvLO05, AGBPM93, APM96, APM01, APM02] but have been simplified for this project because the data collection was carried out through field observation by only one observer. The technological interaction was also analyzed through a simplified range of indicators, aimed mainly at testifying the presence of major actions related to usability: examine/understand the interface, examine/understand the contents (navigation), problems with the interface, problems with the contents, ask help from other visitors, ask help from exhibition staff, look at written instructions, interaction through guide. The last general category of analysis was learning. Some of the widely accepted external indicators for learning include: from an individual point of view, conceptual change, additive knowledge and changes in behaviour; while from the social perspective, verbal interaction (asking questions, explaining the contents, connecting with previous knowledge or the surroundings), collective decision making, conflict resolution, and peer teaching. However, some of them can only be observed either in an experimental or formal learning environment context or through video recording, which allows a more detailed analysis [vLHH02, Pie05]. This is why we chose again a simplified version, adapted to the conditions of the data collection: make comments about the contents, ask/explain contents, connect with previous knowledge, link with surroundings. In any case, the interest of this research was not focused strictly on learning, because the situation did not allow it (goal of the exhibition, conditions of the survey) but on how the different interfaces are used by different kinds of visitors, and it only tried to verify if the most evident indicators of learning were present and why.

3. Qualitative evaluation of ICT applications in Cultural Heritage settings

Having referred to the basic foundations of a methodology for qualitative evaluations of ICT applications in Cultural Heritage settings, we will now discuss some empirical results concerning the use of technological exhibits and how they affect visitors' experience, with particular regard to the social dimension, learning and usability. These come from published findings of related research projects and also from the aforementioned visitors' survey that we participated in at the "Immaginare Roma Antica" exhibition. The statistical analysis remained mainly descriptive because the sample was not big enough to produce reliable results in the non parametric tests and the correspondence analysis. Nevertheless, the findings offer, along with the results of the previous studies, an invaluable repertory of the audience's opinions, attitudes and behaviours in front of different kinds of interfaces and contents, which constitutes a useful basis for future research.

3.1. Social use of exhibits

Even in the case of single visitors, the experience of the exhibition visit consists of a permanent renegotiation between people and objects sharing the same space, in which the resources available in the room, those generated by people through body and verbal language, and visitors' previous experiences are combined. Galani and Chalmers [Gal03, GC03], followed Falk and Dierking's contextual model of the museum visit in their analysis of the influence of the interaction between group members in the relationship established with exhibits in a mixed reality visiting model (one visitor in the gallery and one off-site) at The Lighthouse and the House of an Art Lover in Glasgow. They found out that a shared visit through Hybrid Reality presents some fundamental differences when compared to normal ones: as visual clues cannot be totally activated, visitors organize their visit using the map and, above all, verbal communication, which is primarily devoted to spatial positioning and, secondarily, to talk about the contents, when the latter is usually the main function. This confirms the role played by visual and verbal communication in exhibition visits. With regard to the use of ICT as a museographical tool, the absence of visual clues and the fact of having three different points of view were considered by the authors both as an advantage and a disadvantage: it is an inconvenience because visitors invest a lot of time in agreeing about what they are seeing; but this is also forcing them to pay explicit attention -through description- to the exhibits and explore them in greater depth, as was demonstrated by the fact of linking them to previous knowledge. The conclusion we can derive from this study is that the design of museum experiences has to take intro serious consideration the social dimension because it is so fundamental that visitors will use any available resources to serve it optimally and this might be different than the way it had been originally designed.

Dirk vom Lehn and Christian Heath have also investigated the effect of technology on the social dimension of the visit [vLHH05, vLHH02]. Their main conclusion is that technological exhibits are based on the traditional computer paradigm, which establishes a one-to-one sequential interaction between the user and the machine in order to complete a task and therefore do not allow co-participation and collaboration: there is only one visitor who can use them at a time, while the rest become mere spectators (if they do not know the user) or have to interact through him/her (assisting, requesting navigation paths) if they are visiting together. This was also observed in Rome and was more accentuated in the case of particular interfaces, like a tactile device which forced turn-taking. In the rest of the cases, the size of the screen and the fact that each exhibit was alone in each room, allowed the presence of more visitors and when one user had problems to operate correctly the interface, the other visitors abandoned their passive role and ventured to give advice (this was easier when the user was a child than when he/she was an adult). The same was observed in VR exhibits when they were busy or the observers did not want to use the interface probably due to a lack of skills: then, the visitors did not wait for serious problems to appear but cooperated with the user telling him or her, for example, where to go or what to press in order to carry out certain tasks. Summing up, it appears that cooperation between different groups of visitors was evident in three circumstances: when there were problems with the interface, when the exhibits were busy and when there was a skilled user and the observers did not want to operate the applications themselves. In the case of families or couples, when problems with the navigation appeared, it was not uncommon to see another member of the group abandoning his/her spectator role and trying to help by using the devices at the same time, especially if the first user was a child.

Several studies report that in many cases, when visitors in exhibitions which include ICT applications arrive when the activity has already begun, they get a partial and fragmented experience or the surprise effect intended by designers disappears [vLHH05, HvLO05]. This was also observed in Rome: visitors limited themselves to continuing the exploration where the previous user had left it and missed part of the contents, did not understand how to start the exploration (even having technological skills, like in the case of two skilled teenagers at one multimedia application) and/or repeated what they had seen and did not take the maximum advantage of the interface possibilities (this was more evident in two VR applications, probably because, contrary to other applications, there were no visible instructions). On the other hand, watching what other visitors did, helped those who had little experience with technology to know what to do.

Experimenting with the integration of ICT in exhibitions, many museums initially introduced kiosks and later PDAs as a complementary non-pervasive way of providing supplementary information. However, it has been observed [vLHH05, vLH03] that they both monopolize the visitors' attention, instead of fermenting an egalitarian communication between the different elements involved (user, mobile device, environment, objects and other visitors) as they are designed for a single user. The same was reported by Hsi in the evaluation of the use of PDAs at the Exploratorium in San Francisco [Hsi03]. This obtained results related to ambient, cognitive and attitudinal fields. In the first case, users emphasized a feeling of isolation either with regard to other visitors or with regard to the rest of the exhibition. In the cognitive domain, the researcher discovered that visitors had problems to establish transferences between the real and the virtual world when there were no reference points to allow the superposition of the two kinds of explanation. In all cases, the presence of a virtual guide was a source of motivation and inspiration to try new ways of interaction with the exhibition and to pay more attention to the exhibition discourse. With regard to the attitudinal field, visitors expressed a wide range of interests and preferences concerning contents and presentation format: while some people expected to find exactly the same in the real exhibition and the virtual support, others wanted it to be different or complementary, and this depended on the category of visitor which they belonged to.

Trying to go a step further, some museums have tried to introduce multi-user interfaces, as is the example of a game at the Science Museum in London. But even in this case the system seems to fail because it was designed as a competitive game in which each visitor has to give his/her own opinion instead of coordinating to built shared meanings or complete tasks [vLHH02]: more than collaboration, this supported rather common access to information. On the other hand, if the visitors did not follow all the sequence, they lost track of how their actions were affecting the activity and they augmented their body gestures in order to supply the communication that should be done by the system. It appears that in most cases multi-user interfaces are not yet able to deal with the richness of multi-user collaboration: instead, they either treat different users as a group, that is, again as a single user; or they coordinate internally the inputs preventing the users to understand what was the effect of their intervention. As it was observed in one experimental situation with interactive low-tech displays in which visitors had to collaborate to understand them because the functional parts were physically separated, social interaction and collaboration (through verbal explanations complemented with gestures) were critical for the understanding of the exhibits [vLHH05]: this is why groups had good results while individual visitors had a lot of problems to reach the goal, which indicates that social participation will only exist if it is integrated from early on in the design of the exhibit.

Although the applications at the Rome exhibition were not purposely designed to support group interaction, this does not mean that this did not actually take place. In fact, it was detected in all cases but in some this was the result of problems with the interface. This acted at the same time as problem and as a stimulus. A stimulus because visitors, even strangers, collaborated to understand the operation of the system; but also a problem because this cooperation was focused on the interface and not on the content. This was true even for audiovisuals: in a 3D audiovisual, visitors merely observed the images and only talked when they had problems with the glasses, while in the other exhibit sharing the room (an audiovisual with virtual reconstructions), maybe because there was a voice explaining a story, visitors talked to each other about it. These observations show that the screen paradigm did not help to promote the construction of meaning because it had a "TV effect": activity and people's attention were limited to the observation of the screen, everything was purely visual and thus, in the case of interactive exhibits and as other experimental studies have shown [JRL*98], only the direct user was really involved in the task. This has also a methodological corollary and it is that we would have had the wrong impression if we had only taken into account the timings in each room: as it has also been contrasted in other situations [HvLO05], visitors spend a lot more time in computer-based exhibits than, for example, reading texts, but this is often not due to the fact that the former improve collaborative learning but that in most of the cases they are spending a lot of time to understand -sometimes without success- the interface operation.

Coming again to the Rome exhibition, although in some of the exhibits visitors tried to operate the devices synchronously (in exhibits with two input devices), the only case in which true co-participation was possible was in an Artificial Reality exhibit, because it allowed an interaction not only based on an input device but on the body projected onto a screen. This could be completely demonstrated in the case of a group of elders who had kept a passive attitude in all the rest of the exhibits but here they felt free and comfortable and even found it funny to interact with the e-beings. One of them (a female visitor in her sixties) even took the microphone and explained to her partners the aims of the application. In the rest of the exhibits, as in previous studies [vLHH05], collaboration was limited to assistance in the understanding of the interface operation except in those cases in which they had experience with this kind of environments and they could really concentrate on the exploration of the content.

3.2. Learning

Different studies agree that there cannot be any learning if users do not understand well first the interface operation [vLHH05, Rou04, AGBPM93, SMI03, ASLA*05, EP06, Jov03, XMM05]. In the Rome exhibition this was confirmed in all the cases, even with the audiovisuals: for example, in the aforementioned 3D audiovisual, the only comments made by the observed visitors concerned problems with the glasses. On the other hand, a young couple at a navigable VR application only talked about the content because they did not even need to explore the interface operation, as they were able to navigate automatically. And the same happened in those cases in which a member of the staff acted as a human guide: then people seemed to really enjoy the exhibit and made comments or asked questions about the content. However, in some exhibits, due to their novelty, the technology overshadowed the content: at the tactile device, social exchange was mainly done with the human guide who showed how to manipulate the device but all the comments were about the interface and its utility. Most of the conversations about the content were observed with the multimedia applications, especially if the interface did not pose any operation problems, while in VR it was again comments about navigation, even if it was done through a human guide (people asked what the environment could do and show).

However, when considering learning, we need to take into account the type of visitor and the social interaction that this determines. Because of the learning indicators we had chosen, groups (especially families) gave the impression they learned more because the social interaction between them (e.g. the explanations -sometimes wrong- that, as it has been reported by other studies [FD00, HG94], parents often give to children) provokes more external evidence of learning. In the Rome exhibition, the groups distributed the tasks according to each person's skills: one navigated, the rest looked at the screen or explained the contents. In the case of groups with little knowledge about technology (for example elders), the roles were adopted depending on the personality of each member: usually the one who was more courageous adopted a leading role and explained the contents or even tried to use the application. One group of elders also demonstrated, confirming the constructivist theory, that learning is not only a matter of facility with the interface but also of familiarity with the content: to interest them and/or maintain their attention, the content must appeal to users' previous knowledge and then it is likely that they will try to approach the exhibit, even if they are "afraid" of the technology. In the two cases in which this happened at the Rome exhibition -with two VR applications-, visitors said they wanted to buy the CD-ROM, if it was available, to be able to explore it more carefully at home.

3.3. Interaction with exhibits

After conducting field observations and interviews about operation and learning in several exhibitions, Viviane Jovet [Jov03], who was interested in the study of VR integration in the exhibition as a communicative tool, reached the conclusion that high-tech exhibits have problems of integration in the exhibition because they are a very specific tool; an autonomous and very particular communication system located inside a bigger one that, contrary to the foregoing, is contextual and already owns other completely integrated resources. Furthermore, the visitor has to face two problems: decipher the message and the operation of the intermediary, which usually is not natural or intuitive. This could be seen at Rome exhibition, where, for example a young woman had problems to use the trackball although it might seem very easy. On the contrary, no child had problems to navigate with a Playstation device.

In relation to software interface, it appeared that some people had difficulties exploring one multimedia exhibit when they wanted to find specific objects: this offered too many options and did not have clear "instructions" or a direct way to know how to operate them. On the other hand, another multimedia program, was more intuitive because it had a simpler interface, with only one screen and few buttons. Visitors could navigate this without problems despite the barriers to basic comprehension posed by the language (it was in Spanish). It is also interesting to note what happened with several adult visitors at two different multimedia applications: their first impulse, without having read the instructions, was to click on the visual symbols in the screen because they thought this would get the program started (the application was in English, so they were interpreting the icons); they had not understood that those were the instructions. Seemingly, some of them did not understand that they had to wait for the computer to render the model and kept clicking the buttons again and again thinking it did not work. Finally, as we mentioned above, the fact of starting the exploration at the point were the previous visitor had left it caused problems even to those who had experience with computers. It is important to bear this natural impulses in mind when designing multimedia environments, for example by clearly indicating that the machine is processing, by introducing an automatic "restart" after a time of inactivity and by creating more universally intuitive navigation symbols.

The same applies to VR environments: in one case, the moving books, abstract guides and doors to change level were not evident to the users, so visitors were not using them and consequently were not making good use of all the environment's capacities. VR designers have to remember that not all users will naturally have the impulse to explore in detail a virtual environment to discover its possibilities, especially in the exhibition context; visitors need instructions or, even better, signs clearly indicating their function. Another issue is how to display knowledge uncertainty in virtual models: one of the most frequently used conventions is to leave the corresponding element with a uniform color. However, this is not always understood by non-specialists. One family wanting to see the appearance of one artistic monument at a specific date was surprised to see that "it was all white!", not having realized that this colour was indicating the lack of knowledge for that specific period. VR has also problems of navigation, especially with avatars: the problems some visitors encountered in one avatar-navigable application (not knowing how to make the avatar approach specific elements, or what it was looking at and what actions it had unchained) showed that, as other authors have demonstrated [Sch97], it is more intuitive to navigate in first than in third person. Even the most intuitive interfaces have the problem of not being natural enough. The Artificial Reality application allowed corporal direct interaction with the e-beings but still required the use of a microphone: in general, people had not understood that the beings learned words and were getting disappointed when they were asking them questions and did not get any answers back; in another case, one old lady was talking very quietly close to the screen without using the microphone. In other cases, as visitors could only see the microphone, they thought the only interaction they were allowed was talking and did not try to move in front of the screen. This shows that intuitiveness demands, as an essential feature, visibility because most of the times people will not read written instructions and will interact directly with the exhibit; therefore, it has to make all its capabilities evident at first sight.

Heath and vom Lehn demonstrated that even the most evident and basic levels of interactive exhibits can cause problems, for example, depending on the way visitors are approaching the resource [HvL02]. This applies also to high-tech displays, as the case of one young couple in the Rome exhibition demonstrates. Even though there was a rope clearly separating the two spaces corresponding to each of the two screens, and the two images were different, users tried to watch the second application while sitting on the first and made comments about the bad quality of the visualization, as they had not understood that each screen was a different application and that the polarized glasses could only be used in the 3D hyper-realistic audiovisual. Similar findings were recorded by a group of Spanish sociologists [ASLA*05] who conducted a study about the coexistence of technological and traditional exhibits in a temporary exhibition about the Iron industry in the Basque Country specifically designed for research purposes. This demonstrated that people had problems establishing a relationship between the real object and its virtual reconstruction despite the obvious association created by the fact that they were in the same exhibit, a combination of virtual workbench and showcase. The results also showed problems of integration of this Virtual Showcase inside the exhibition context: it became the main attraction of the exhibition, and this generated great expectations which, in most of the cases, could not be satisfied.

From the observation at the Rome exhibition it has been possible to group visitors' interaction in four large categories. The first category is that of passive exploration of the exhibition. In several cases, visitors seemed to expect to see audiovisuals and were not looking for direct interaction with the exhibits. Some people preferred to see, to receive or be given (visual) information, than to interact; or if they did use the application, they only looked for (navigable) images/reconstructions or audiovisuals and skipped any textual explanations. This is why instead of concentrating on the content, they often asked the guide to show them the visualization possibilities of the environment. In fact, as the general evaluation showed [FPP06], most of the visitors came to see reconstructions of Rome and associated VR with computer-modeled images of buildings. The second type of visitors were couples. In most cases, especially if the couple was not young, it was the man who controlled the interface while the woman adopted a passive role. In the case of young couples, the girl also tried to use the application but her boyfriend would finally take the control of the interface because he usually seemed to have more facility using it. This was a good example to compare the differences in navigation style between adults and children: the first look at the interface and try to infer how it works before starting, while children are more impulsive and have a more hazardous pattern, clicking at different points or pressing all the buttons at the same time in order to get the system to respond.

The third type of visitors were families (parents or grandparents with children). Adults would usually let the children interact with the exhibit directly, even if they belonged to different groups of visitors, and they would always act as guides for the navigation and/or the understanding of the content. Consequently, they always filtered the experience through their own perceptions and, in the worse case, gave wrong information about the interface and/or the content. However, sometimes the situation was inverted: twice at the tactile exhibit and once at the VR avatar-navigable application, the father tried first, then the children and finally the mother; when it was her turn, the children explained to her how it worked. If adults did not know how to operate the interface they would ask other visitors or the museum staff (in the case of women) or read the instructions. The observation showed that in order to be able to correctly manipulate the interface and enjoy/learn the content of the various applications, users needed to have a certain age and/or a high level of experience with computers. The last type were older visitors. These showed limited interaction with interfaces, usually adopted a passive role, observing what others did, even when there was opportunity to interact with the application with the help of a guide. In the Rome exhibition older visitors felt comfortable with two exhibits. The first of them was the Artificial Reality application, probably because they perceived it as a game and also because it allowed the possibility of a direct interaction. The second was a touch screen, which had a very easy interface. Despite this grouping into types, there was never a single pattern of navigation, but as other authors have shown [HvLO05] each category might have been different at the same time and could change diachronically according to the composition of the group, the problems encountered during interaction with the exhibit and personal skills or interests.

An interesting conclusion from the observation was that people like realism and are willing to feel immersed in the virtual reconstruction. This is why visitors were making very positive comments about the virtual reconstruction of "Domus Aurea". In one case a young man using a multimedia application with virtual reconstructions was following a navigation path and was moving his body exactly the same way he would in the real environment: gazing at the room/view from the entrance, from left to write and from the eyes' level to the ceiling/sky. In another case, a child had problems with the tactile interface but when he tried the polarized glasses, he immediately extended his arms to touch the three-dimensional virtual image that had suddenly appeared in front of his eyes. Another lesson learned from the observation is that, in general, people seemed to appreciate and prefer a human guide to show them the interface -and then they were able to concentrate on the contents- or assist them in case of problems with the navigation because it was easy for them to reproduce the actions they had already seen or been told. In some cases they did not even try the interface on their own: if there was a member of staff present, they asked him or her directly and then followed the instructions. This is linked with the fact that visitors did not spend a long time exploring the interface operation -as it has generally been reported by different authors [AGBPM93, Alc92]. If they had problems with it, they would in some cases read the instructions (if available) or ask for help but in most of the cases, they explored it at random until they could see some content and abandon the exhibit without having fully exploited its capacities. This behavior was observed especially in multimedia applications, while in VR ones users spent more time because they had one immediate result: the "physical" motion inside the environment.

4. Conclusions

Constructivist theories support the idea that learning is produced through interaction with the content inside a socially dynamic environment in which physical activities and later language is fundamental in order to build shared meanings and integrate knowledge. As this is especially true in the informal learning environments, it should be taken into consideration by curators and designers when introducing ICT exhibits in museums. The problem is that they often overlook the fact that ICT's interaction model is individual, oriented to the achievement of a goal through a sequence of strictly defined actions and responses [HvLO05]. This is what makes them so effective for individual learning but also so different from the processes that arise in the social and rich in resources exhibition context, where their integration as a communication tool is often problematic [Puj05, EP06].

However, several ways are opened that can offer different solutions. One possibility is to keep the interaction paradigm and then to use ICT to complement the experience of the visit. As the major evaluation of the permanent exhibition of the Holocaust Memorial Museum in the United States has shown, this can be done in four ways [Swi05]: by adding the interactive, personalized layer in travelling exhibits, which are necessarily "flat"; through the Internet, by preparing or complementing the visit and creating a forum of debate or even a community of people related to the museum; by promoting long term learning thanks to reiterate remote access; and finally, in all cases, by offering the possibility of personalization which, contrary to traditional mass media, ICT are able to provide. The other possibility is to change the interaction paradigm. If even the multi-user environments isolate the users and/or create problems for them coping with non-sequential collaboration, we can try to use real objects as interface. This is exactly the goal of "Tangible User Interfaces", which are starting to be tested in several projects [XMM05, SFD03, PPW05] and offer the advantage of requiring little time from the user to learn how to use them because they rely on our way of interacting with the physical world and they support the complexity of the multi-layered collaboration between several users [XMM05].

All cases and situations demonstrate that visitors coming in groups adopt a social modality of visit which predisposes them to interact simultaneously with the exhibits. But in most cases the interfaces do not allow it and therefore they have to adapt their behavior depending on three major categories of factors: the exhibit (kind of interface and interest of contents), the environment (if busy or not) and visitors (composition of the group, personality and skills). In order to discriminate and understand the significance of each factor, we need more integrated evaluations conducted in optimal conditions, moving beyond attitude and usability to include more complex phenomena such as learning and social use of exhibits. We also need to continue developing specific tools for observational analysis of technological displays in museums: detect observational indicators, make hypothesis about causal factors, and then test them experimentally and on site. To that end, video recording could be specially useful [vLHH02, Pie05]. This can be completed with other traditional evaluation systems (interviews, concept mapping, interaction logging, focus group discussions) in order to obtain a deeper and wider insight.

5. Acknowledgements

The findings from the "Building Virtual Rome" exhibition were obtained during a collaborative project between the authors and the Istituto per le Technologie Applicate ai Beni Culturali of the CNR in Rome. We would like to thank Dr Maurizio Forte and Dr Sofia Pescarin for their kind invitation to participate in this experience. The work presented in this paper is part of the wider research network 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.

References

[AGBPM93] ASENSIO M., GARCÍA BLANCO A., POL MÉNDEZ E.: Evaluación cognitiva de la exposición "los bronces romanos en españa": dimensiones ambientales, comunicativas y comprensivas. *Boletín de ANABAD 43*, 3-4 (1993), 215–255.

[Alc92] ALCALDE G.: *La Difusió de l'arqueologia mitjançant els museus arqueològics, : avaluació dels visitants dels museus arqueològics de Catalunya i anàlisi dels conceptes que aquests museus transmeten al públic.* Tesi doctoral, Universitat de Girona, 1992.

[APM96] ASENSIO M., POL MÉNDEZ E.: Cuando la mente va al museo: un enfoque cognitivo-receptivo de los estudios de público. In *IX Jornadas estatales DEAC-Museos: La exposición*. Diputación provincial, 1996, pp. pp. 83–133.

[APM01] ASENSIO M., POL MÉNDEZ E.: *Planificación en museología: el caso del Museu Marítim de Barcelona*. Manuals Museu Marítim. Museu Marítim, Barcelona, 2001.

[APM02] ASENSIO M., POL MÉNDEZ E.: Los cambios recientes en la consideración de los estudios de público: la evaluación del museu d'història de la ciutat de barcelona. In *II Congrés internacional sobre museïtzació de jaciments arqueològics. Nous conceptes i estratègies de gestió i comunicació*. Museu d'Història de la Ciutat, Barcelona, 2002, pp. pp. 310–321.

[ASLA*05] ALZÚA-SORZÁBAL A., LINAZA M. T., ABAD M., ARRETXEA L., SUSPERREGUI A.: Interface evaluation for cultural heritage applications: the case of ferrum exhibition. In *VAST 2005: 6th International Symposium on Virtual reality, Archaeology and Intelligent Cultural Heritage*, Mudge M., Ryan N., Scopigno R., (Eds.). Eurographics Ass., Pisa, 2005, pp. pp. 122–128.

[EP06] ECONOMOU M., PUJOL L.: 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., Kalay Y., (Eds.). The Faculty of Architecture, University of Hong-Kong, Hong Kong, 2006, pp. pp. 284–302.

[FD92] FALK J. H., DIERKING L. D.: *The Museum Experience*. Whalesback Books, Washington, D.C., 1992.

[FD00] FALK J. H., DIERKING L. D.: *Learning from Museums. Visitor Experiences and the Making of Meaning*. AltaMira Press, Oxford, 2000.

[FPP06] FORTE M., PESCARIN S., PUJOL L.: Vr applications, new devices and museums. In *Proc. 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2006) Projects and Short papers volume*. Eurographics, Nicosia, Cyprus, 2006.

[Gal03] GALANI A.: Mixed reality museum visits: using new technologies to support co-visiting for local and remote visitors. In *Museological Review*, vol. Special issue 10 (spring). Department of Museum Studies, Leicester, 2003, pp. 1–17.

[GC03] GALANI A., CHALMERS M.: Far away is close at hand: shared mixed reality museum experiences for local and remote museum companions. In *ICHIM03: Cultural institutions and digital technology*. Ecole du Louvre, Paris, 2003.

[Hei82] HEIN G. E.: Evaluation of museum programmes and exhibits. In *Museum Education*, Hansen T. H., Andersen K. E., Vestergaard P., (Eds.). Danish ICOM/CECA, Copenhagen, 1982, pp. 21–26.

[HG94] HOOPER-GREENHILL E.: *Museums and Their Visitors*. Routledge, London, 1994.

[Hsi03] HSI H.: A study of user experiences mediated by nomadic web content in a museum. *Journal of Computer Assisted Learning 19* (2003), 308–379.

[HvL02] HEATH C., VOM LEHN D.: Misconstruing interactivity. In *Interactive Learning in Museums of Art and Design*. London, 2002.

[HvLO05] HEATH C., VOM LEHN D., OSBORNE J.: Interaction and interactives: collaboration and participation with computer-based exhibits. *Public Understanding of Science 14* (2005), 91–101.

[Jov03] JOVET V.: 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*. Ecole du Louvre, Paris, 2003.

[JRL*98] JOHNSON A. E., ROUSSOU M., LEIGH J., VASILAKIS C. A., BARNES C. R., MOHER T. G.: The nice project: learning together in a virtual world. In *Proceedings VRAIS'98*. 1998, pp. pp. 176–183.

[Mac93] MACDONALD S.: The enigma of the visitor sphinx. In *Museum Visitor Studies in the 90s*, Bicknell S., Farmelo G., (Eds.). Science Museum, London, 1993, pp. pp. 77–81.

[Man01] MANTOVANI F.: Vr learning: potential and challenges for the use of 3d environments in education and training. In *Towards CyberPsychology: Mind, Cognitions and Society in the Internet Age*, Riva G., Galimberti C., (Eds.). IOS Press, Amsterdam, 2001, pp. pp. 207–226.

[McM88] MCMANUS P. M.: Good companions ... more on the social determination of learning-related behaviour in a science museum. *International Journal of Museum Management and Curatorship 7* (1988), 37–44.

[MW05] McCarthy J., Wright P.: Time, place and technology in museums: a dialogical approach to the experience. In *Re-thinking Technolgy in Museums: Towards a New Understanding of People's Experience in Museums*, Ciolfi L., Cooke M., Bannon L., Hall T., (Eds.). Limerick, Ireland, 2005.

[NG02] Ng K. H.: *Mixed Reality interaction for group experience*. Master of research dissertation, The University of Nottingham, 2002.

[OBP05] Owen R., Buhalis D., Pletinckx D.: 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., Scopigno R., (Eds.). Eurographics Ass., Pisa, 2005, pp. pp. 129–136.

[Pie05] Pierroux P.: The language of contextualism and essentialism in museum education. In *Re-thinking Technolgy in Museums: Towards a New Understanding of People's Experience in Museums*, Ciolfi L., Cooke M., Bannon L., Hall T., (Eds.). Limerick, Ireland, 2005.

[Pod04] Podgorny J.: *Studying visitor engagement in Virtual Reality based children's science museum exhibits*. Master of arts program in the social sciences, The University of Chicago, 2004.

[PPW05] Petridis P., Pletinckx D., White M.: A multimodal interface for presenting and handling heritage artifacts. In *Virtual reality at Work in the 21st Century. Impact on society*, Thwaites H., (Ed.). International Society on Virtual Systems and Multimedia, 2005, pp. pp. 91–100.

[Puj05] Pujol L.: Interactivity in virtual and multimedia environments: a meeting point for education and ict in archaeological museums. In *Virtual reality at Work in the 21st Century. Impact on society*, Thwaites H., (Ed.). International Society on Virtual Systems and Multimedia, 2005, pp. pp. 37–52.

[Rou04] Roussou M.: Interactivity and conceptual learning in virtual environments for children. In *CHI 2004. Doctoral Consortium*. Vienna, 2004, pp. pp. 1049–1050.

[Sch97] Schroeder R.: Networked worlds: social aspects of multi-user virtual reality technology. *Sociological Research Online 2* (1997), http://www.socresonline.org.uk/socresonline/2/4/5.html.

[SFD03] Strauss W., Fleischmann M., Denzinger J.: Knowledge spaces - experience based learning and cultural education in the media age. In *ICHIM03: Cultural institutions and digital technology*. Ecole du Louvre, Paris, 2003.

[SMI03] Scott S. D., Mandryk R. L., Inkpen K. M.: Understanding children's collaborative interactions in shared environments. *Journal of Computer Assisted Learning 19* (2003), 220–228.

[Swi05] Swiader L.: Using technology in the permanent exhibit of the united states holocaust memorial museum. thoughts after a major evaluation. In *Re-thinking Technolgy in Museums: Towards a New Understanding of People's Experience in Museums*, Ciolfi L., Cooke M., Bannon L., Hall T., (Eds.). Limerick, Ireland, 2005.

[VKT*01] Vlahakis V., Karigiannis J., Tsotros M., Gounaris M., Almeida L., Stricker D., Gleue T., Christou I. T., Carlucci R., Ioannidis N.: *ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites*. Tech. rep., 2001.

[vLH03] vom Lehn D., Heath C.: Displacing the object: mobile technologies and interpretive resources. In *ICHIM03: Cultural institutions and digital technology*. Ecole du Louvre, Paris, 2003.

[vLHH02] vom Lehn D., Heath C., Hindmarsh J.: Video based field studies in museums and galleries. *Visitor Studies Today! Fall/Winter 2002* (2002), 15–23.

[vLHH05] vom Lehn D., Heath C., Hindmarsh J.: Rethinking interactivity: design for participation in museums and galleries. In *Re-thinking Technolgy in Museums: Towards a New Understanding of People's Experience in Museums*, Ciolfi L., Cooke M., Bannon L., Hall T., (Eds.). Limerick, Ireland, 2005.

[XMM05] Xu D., Mazzone E., MacFarlane S.: Informant design with children. designing children's tangible technology. In *Re-thinking Technolgy in Museums: Towards a New Understanding of People's Experience in Museums*, Ciolfi L., Cooke M., Bannon L., Hall T., (Eds.). Limerick, Ireland, 2005.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Destroying Cultural Heritage: Technical, Emotional and Exhibition Aspects in Simulating Earthquake Effects on a Gothic Cathedral

L. Chittaro†, R. Ranon and D. Corvaglia

HCI Lab, University of Udine, Italy

Abstract

While a significant research effort has been devoted to produce virtual reconstructions of cultural heritage, the issue of reproducing the effects of natural or man-provoked disasters (e.g., earthquakes, floods, wars) on cultural heritage has received much less attention. Moreover, presenting these events with multimedia installations on museums requires to consider how to properly convey the dramatic aspects of the experience besides the faithful simulation of the damage caused. In this paper, we focus specifically on earthquakes and their effects on historical buildings. We present the methodology we have followed to produce a museum experience of a real earthquake that struck a gothic cathedral. We discuss technical (e.g., building a 3D model that is suitable to the considered purpose), emotional (e.g., testing the exhibit with pilot studies on users), and exhibition aspects (e.g., using infrasound to increase the realism of the experience and the dramatic feelings it evokes).

1. Introduction

While a significant research effort has been devoted to produce virtual reconstructions of cultural heritage (e.g., [GCR01, MVSL05, STY*03, dHCUCT04]), reproducing the effects of natural or man-provoked disasters (e.g., earthquakes, floods, wars) on cultural heritage has received much less attention. In the latter case, there are new technical, exhibition and methodological issues that need to be considered.

From a technical point of view, cultural heritage objects need to be modeled and rendered in such a way that the effects of disasters can be properly visualized. In the case of buildings, for example, 3D modeling and rendering must take into account damages to structural elements (e.g., broken walls). When one wants to realistically simulate the effects of past disasters, an additional issue concerns how to effectively combine and exploit the available sources of information (e.g., photographs of the objects after the disaster) or to compensate for the lack of data (e.g., sufficiently detailed structural and construction data are unlikely to be available for old buildings).

In presenting the results of modeling and (possibly) simulation in the context of multimedia experiences, e.g., installations in museums, emotional and exhibition aspects play a major role. For example, how can we properly convey the dramatic feelings of the experience besides the faithful visualization of the damage caused?

Finally, there is also the problem of finding a proper design methodology and identifying necessary sources of information and technical skills needed in producing this kind of virtual experiences.

In this paper, we focus on earthquakes and their effects on historical buildings, and present a case study involving the production of a museum experience concerning a 6.4 Richter magnitude earthquake that on 1976 hit the Friuli region in Italy causing about 1000 casualties and severe damage to the historical and cultural heritage of the region. The museum experience concentrates on what happened to a symbol of the Friuli region, i.e. the gothic cathedral in the town of Venzone (which now has been fully rebuilt). The project has two main goals:

- to obtain a realistic physically-based animation (through simulation) of the effects of the earthquake on a 3D model the cathedral;

† chittaro@dimi.uniud.it

- to present the rendered animation in a museum multimedia experience able to convey emotions and feelings related to the event.

Although our case study refers to a specific kind of disaster and cultural heritage object, we identify in the paper a number of general issues and possible solutions that can be applied in related projects. Moreover, we mostly use off-the-shelf and relatively affordable hardware and software, so that the ideas in the paper can be easily adopted by teams that want to develop computer-generated experiences of the effects of disasters on cultural heritage.

The paper is structured as follows. In section 2, we outline the general design process that has been followed. Section 3 provides more detail on the technical aspects of the 3D modeling and simulation of the earthquake effects on the cathedral. In Section 4, we describe how the museum experience was designed, including details on how to increase the realism of the experience and the dramatic feelings it evokes using infrasound. Section 5 concludes the paper and outlines future work.

2. Methodology

In the first weeks of the project, we interviewed domain experts (i.e., seismic engineers, civil engineers and experts in the analysis of earthquake effects on churches) and witnesses. By witnesses, we mean people who personally experienced the considered earthquake in Venzone, since there were no witnesses of the cathedral collapse. Moreover, we collected relevant sources of information: photographs and drawings of the cathedral before and after the earthquake, relevant newspaper articles, books, and film footage.

The goal of interviewing domain experts was to identify an effective way of simulating the effects of the earthquake on the cathedral. Ideally, with detailed data about the structural and mechanical features of the building, as well as about the forces applied by the disaster, reproducing the effects is mainly a matter of finding an effective simulation technique. In practice, we found that available data were insufficient for this approach, and we believe that this issue is not peculiar to our project: it is often hard to collect detailed structural information about cultural heritage objects, and it can be even more difficult to have accurate data about the physics of a disaster in the precise location of the cultural heritage object. On the other hand, it is easier to get information about the state of the cultural object after the disaster (in our case, photographs of the cathedral taken after the earthquake). However, these sources of information cannot be directly used as inputs to simulation, but they rather can be used to evaluate how realistic it is.

After interviewing the experts, we chose to follow a methodology based on starting with a coarse visual simulation (using parameters given by seismic engineers), and successively refining it by having different kind of domain experts evaluating the results (also matching them with the photographs of the cathedral after the earthquake). The main advantage of this approach is that it was much easier for the experts, by visually evaluating the simulation result, to elicit and refine their knowledge and hypotheses about the earthquake effects on the cathedral (and then suggest refinements to the simulation). This allowed us to properly take into account structural modifications or consolidation works that were known to have been performed on the cathedral in the past, but about which the experts could initially only make vague hypotheses with respect to simulation. As we will see in Section 3, special 3D modeling operations were performed to ease the work of the experts in evaluating the accuracy of the simulation and suggesting improvements.

In a second phase, we identified a group of pilot users to test prototypes of the museum experience. The user group was composed by witnesses (to test the fidelity of the experience from a perceptual and emotional point of view) and people that had no previous experience of major earthquakes (to check if the lack of episodic memory could play a role and make the experience less intense).

In the interviews with witnesses, the main purpose was to identify important aspects in their emotional and sensorial experience of the earthquake, such as sounds, images, events and feelings that they felt to characterize their experience. Two main topics recurred in most interviews. First, sound played a very important role. Therefore, we decided to focus part of the museum experience on this aspect, and to investigate how to effectively render a proper earthquake aural experience (see Section 4.2). Second, many witnesses' memories highlighted that clouds of dust made it difficult to see a few seconds after the earthquake start. Faithfully reproducing the latter aspect in the simulation would have contrasted with the goal of visualizing how the cathedral collapsed, so we decided to limit visual occlusion caused by clouds of dust.

The work was then divided into three main parts: 3D modeling and simulation, sound production, and museum experience design. In each part, we used an iterative design process, with prototypes (simulation renderings, sounds, storyboards) evaluated by the whole team, by experts, and with pilot users.

3. Modeling and Simulation

In modeling cultural heritage objects for simulating and visualizing the effects of disasters, we followed an iterative process that can be decomposed into the following main steps:

- first, build a 3D model where all parts that have a role in the simulation (e.g., structural elements of the cathedral) are modeled as separate objects. In this step, the 3D model can be composed by just geometries, without any shading information. When objects break into parts because of the

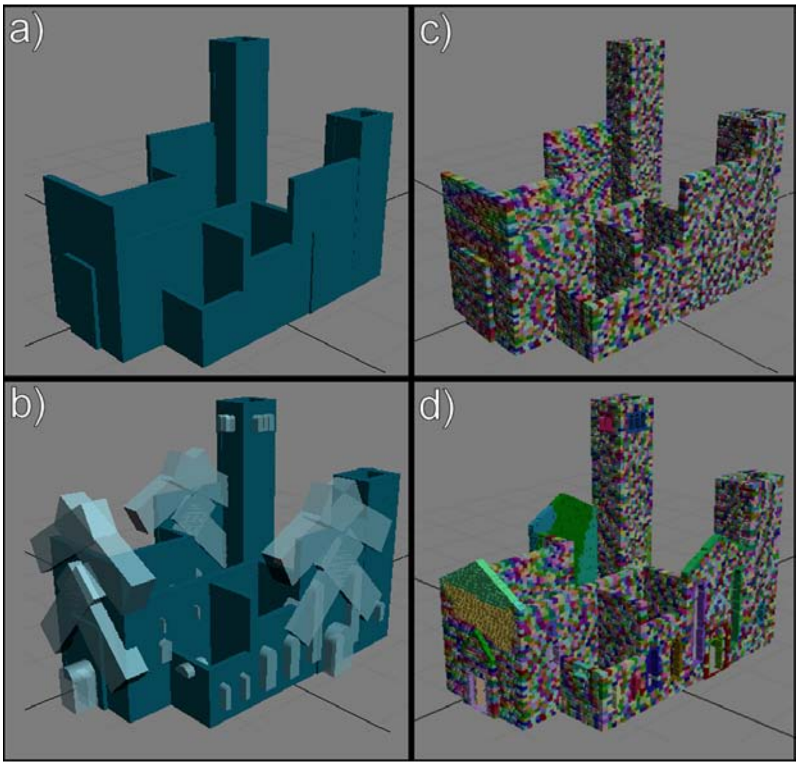


Figure 1: *(a) coarse 3D model of the cathedral using just bounding boxes of the cathedral walls; (b) the same model with geometries to be subtracted visualized with semi-transparent color; (c) detailed model with single stones ; (d) detailed model with single stones after the volume subtraction operations.*

disaster, we model those parts as separate objects, exploiting information about the object after the disaster (in particular, photographs of the cathedral after the earthquake) to infer which parts needed to be modeled;

- second, enter physics-related information about the various modeled parts into the simulation, together with the available data about the disaster (forces applied to the building by an earthquake shock) and run the simulation;
- third, apply the results of the simulation to the 3D model (deriving an animation), and test it with domain experts and against the available sources of information (in particular, photographs taken after the disaster). When the result is not satisfying, either the 3D model needs to be improved (e.g., by modeling new parts), or the simulation parameters need to be refined;
- fourth, model additional objects that have no role in the simulation, add effects like dust clouds and fragments, as well as lights and shading information. The final result is then rendered to a movie or possibly converted into a format suitable for real-time interaction.

An interesting alternative that has the potential to greatly simplify the modeling step is to employ techniques that are able to automatically (and in a visually convincing way) compute breaks, cracks or tears [OH00, OH99, PKA*05]. However, some of these methods are mostly intended to ease the work of the animator by generating physically plausible animations, but have not been demonstrated to be able to produce physically faithful animations [OH00] (and this is a requirement in our case study). Other methods [PKA*05] are too complex for simulating an entire building. Finally, none of these methods is, to the best of our knowledge, integrated into tools that can be readily used for production.

In the following, we explain how these steps have been carried out in our case study, particularly focusing on the first three ones.

3.1. Modeling

The entire modeling and simulation steps have been carried out using 3DS Max 6 and the Reactor plug-in for physics-based simulation.

First, we created a coarse model of the cathedral by using bounding boxes of the cathedral walls (see Figure 1a), and a number of volumes (to be subtracted from the walls) to model holes (doors, windows), archways and the roof profile (see Figure 1b).

Single stones in walls were then automatically created by scripts we wrote to the purpose (the scripts allow one to fill a bounding box with stones with different patterns, randomly determine the size of each stone inside a given interval). Figure 1c shows the application of the scripts to the initial boxes model, while Figure 1d shows the final result after the subtraction operations that carve holes and modify the upper contour of the walls. At the end, the fully detailed cathedral model contained about 15000 stones.

The last step before simulation consisted in coloring the stones on the basis of their movements caused by the shock, to help the experts in evaluating the outcome of the simulation. Starting from photographs of damages to the cathedral taken after the earthquake, experts drew lines over them to highlight so-called macro-elements, i.e., portions of walls (divided by so-called fracture lines) that "stayed together" at least in the initial instants of the earthquake. For example, Figure 2 shows one of the photographs with the superimposed lines drawn by the experts, and the resulting colors that were applied to the 3D model of that part of the cathedral.

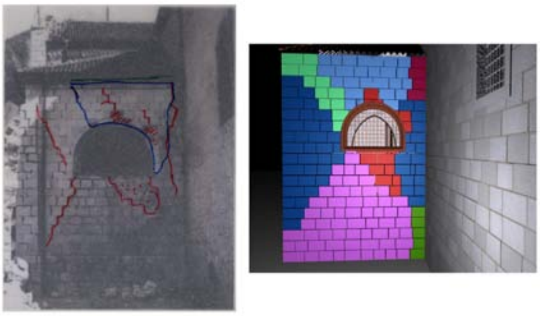


Figure 2: *(left) photograph of part of the cathedral after the earthquake with lines highlighting relevant macro-elements; (right) resulting coloring of cathedral stones (red indicates stones that fell down after the shock, color changes highlight fracture lines).*

At this point, the boxes and the colored-stones model were fed into the simulation. The 3D model was successively refined by adding detailed models of architectural elements, such as windows, portals, and statues, and shading information (lights, materials, textures). The final 3D model of the cathedral can be seen in Figure 3.

3.2. Simulation

As mentioned previously, accurate simulation of the effects of the earthquake on the cathedral would require:

- to know how the terrain under the cathedral moved because of the shock. This would in turn require both to have sensors (e.g., accelerometer) recordings of the movement under the cathedral, as well as to model the mechanical features of the terrain below it. Neither of the two information was available.
- to have an accurate structural and mechanical model of the cathedral. This is problematic for old buildings, since they typically do not have construction plans available and, furthermore, several modifications or consolidation operations may have been done on them over centuries (and this was the case for the building under consideration), thus making it very difficult to predict mechanical behavior.

Figure 3: *The final 3D model of the cathedral.*

As explained in Section 2, the strategy we used to compensate for the lack of sufficiently detailed data was to start with a coarse simulation, and then progressively refine it together with domain experts. In the following, we explain in detail the adopted simulation procedure.

The initial coarse simulation was done on the bounding box model of the cathedral, using as input the 3D movements derived by a sensor recording of the 1976 earthquake measured by accelerometers at a nearby dam. As shown in the plot in Figure 4, the recording describes terrain movements in time along the North-South, East-West and vertical axes. Domain experts suggested values of parameters such as wall mass, elasticity and strength of structural linkages between walls, and corrections to be done on the movement recording (to take into account the distance between the cathedral and the dam) and terrain differences (the dam accelerometers were mounted on harder, rocky terrain). The goal of this initial step was to obtain, taking into account the whole architectural structure of the cathedral, the movements in time of each wall determined by the shock.

After that, the simulation was divided into parts, with each part dedicated to an independent zone of the cathedral. Nine

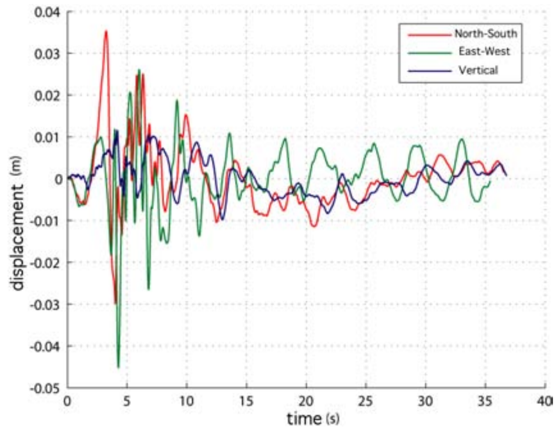


Figure 4: *Plot of the earthquake shock as measured by accelerometers in a dam a few kilometers from the cathedral.*

different zones were determined by the experts, considering the photographs taken after the earthquake (such the one in the left of figure 2). Each zone presented damages that had no causal relations (e.g., due to collisions between stones) with damages in other zones. Each zone simulation used only the relevant part of the 3D colored-stones model, with mechanical parameters (mass, elasticity, friction) defined for each stone, and mechanical links (such as springs) to model physical linkages between stones. Moreover, each simulation used as input the wall movements computed in the initial global simulation. This allows to take into account the whole structure of the building while we are computing simulation on just a part of it. Moreover, by dividing the simulation into local zones, we were able to reduce the time to compute each simulation, thus making the refinement process easier (each zone simulation took from a few minutes to a few hours, depending on the number of structural elements that were part of the zone). Each local simulation was iteratively refined by examining the results with the domain experts, who then suggested modifications on the basis of their knowledge about the cathedral and their hypotheses about its structural properties. For example, Figure 5 shows some frames from one of the computed local simulations which refers to the south facade of the cathedral. Blue-colored brick stones stayed in place after the shock, while purple stones are a fracture zone that fell down to the ground, successively breaking into pieces because of the reached speed.

Additional structural constraints were often introduced to model known structural modifications (e.g., consolidation operations) that in the past were carried out on the cathedral. In particular, this was done by introducing more or less rigid links in locations suggested by the experts. For example, the left part of Figure 6 shows the west facade of the cathedral; note that the upper part of the facade falls to the ground but does not break into single pieces: in the real cathedral, this was due to injections of concrete that were performed some years before the earthquake.

Once experts thought the simulation had reached a satisfactory level of realism, we integrated the local animations together, deriving a full animation of the effects of the earthquake on the cathedral. Dust clouds and stone fragments were then added by using particle systems, as can be seen in the right part of Figure 6.

4. Museum Experience

In this Section, we describe how the museum experience has been designed and developed. The experience is based on a short movie †, which is shown to museum visitors inside a properly equipped projection room. In the following, we briefly describe what the movie shows, provide more details the work done on sounds, and on the projection room.

4.1. The Movie

The movie combines the computer-generated animation with existing film footage to recreate the experience of the earthquake. It is structured in three parts.

The first part, using pre-earthquake film footage, shows the town surrounding the cathedral. Its purpose is to introduce the viewer to the experience and provide an historical context.

The second part is rendered from the simulation described in the previous Section. The transition from the first to the second part uses a wireframe model of the cathedral (which is then shaded after a few seconds) to convey the idea that the viewer is going to see a computer-generated simulation. After a flyby showing the cathedral from various points of view, the earthquake begins. The simulation part lasts about one minute (as the 1976 earthquake did) and exploits camera movement techniques that are inspired by disaster movie scenes to increase the emotional involvement of the viewer.

The third part exploits film footage to show the enormous reconstruction effort that took place in the area (the earthquake killed about one thousand people and destroyed or damaged thousands of buildings). The movie ends with ordinary scenes from today's life, to relax the viewer and show the successful result of the reconstruction of both the cathedral and its surrounding area.

4.2. Modeling Sounds

There are multiple sound sources that combine to create an earthquake aural experience: the typical rumble of the

† the movie can be watched on the Web site dedicated to the project: http://hcilab.uniud.it/earthquake/

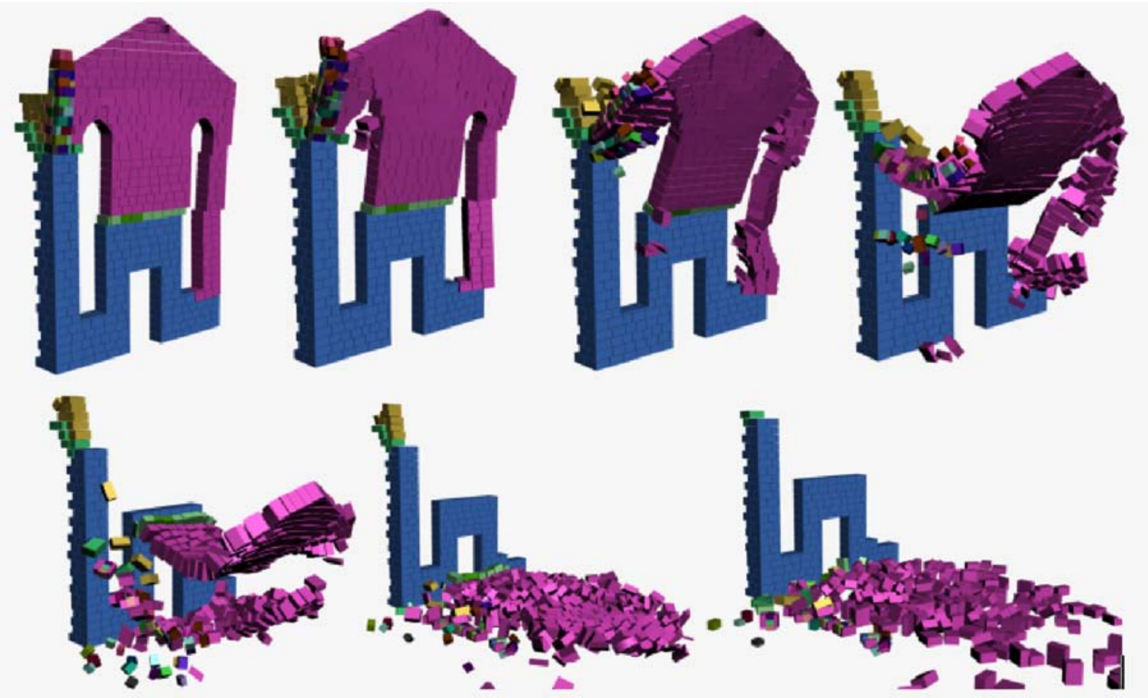


Figure 5: *Some frames from the (local) simulation of the south facade of the cathedral .*

earthquake, nearby buildings falling down, glasses breaking, landslides on close mountains, etc. These sound sources have different spatial locations, generally fill the auditory channel of the listener, and have a primary role in generating emotions of fear. Moreover, an earthquake produces infrasonic frequencies (i.e., below 20 Hz) that cannot be heard, but that are perceived through the body and have a role in producing feelings of panic, fear and sense of disorientation [OCA04].

Since we did not have any accurate audio recordings from the 1976 earthquake, the work on sounds was divided into two main activities: producing and combining the above mentioned different sound sources and finding effective ways to use infrasound and deliver it to visitors.

With respect to the first activity, we started by plotting in the frequency domain (see Figure 7) the movements produced by the 1976 earthquake as measured by the available sensors (plotted in Figure 4). The idea was to use the result to shape the earthquake typical rumble.

Most sound sources in the movie were derived from recordings of seismic events (earthquakes and tsunamis), or of related events (such as stones falling to the ground, landslides). These sounds were then filtered and assembled together to recreate a plausible aural experience, also according to the corresponding events shown in the movie, and by using positional audio for some of them (e.g., for stones that roll towards the viewer).

With respect to the second activity, we used two kind of infrasound. In the movie, before the earthquake starts, we use a continuous infrasound (with a frequency of 12 Hz), produced by using a software low-frequency generator, to induce changes in the emotional state of viewers [OCA04]. During the earthquake sequences, infrasound taken from recordings of real earthquakes and tsunamis are used.

The audio part was encoded according to the Dolby Digital 5.1 standard, which is the de-facto solution in the movie industry ‡.

4.3. The Museum Room

The museum room dedicated to the earthquake experience is based on the following components:

- DVD player with optical outs;
- DLP projector;
- Audio decoder able to reproduce infrasound. Typical consumer audio decoders cut very low audio frequencies (below 20 Hz), and as a result do not allow the reproduc-

‡ the movie available on the Web site uses stereo sound and does not contain infrasound for ease of playback on standard equipment

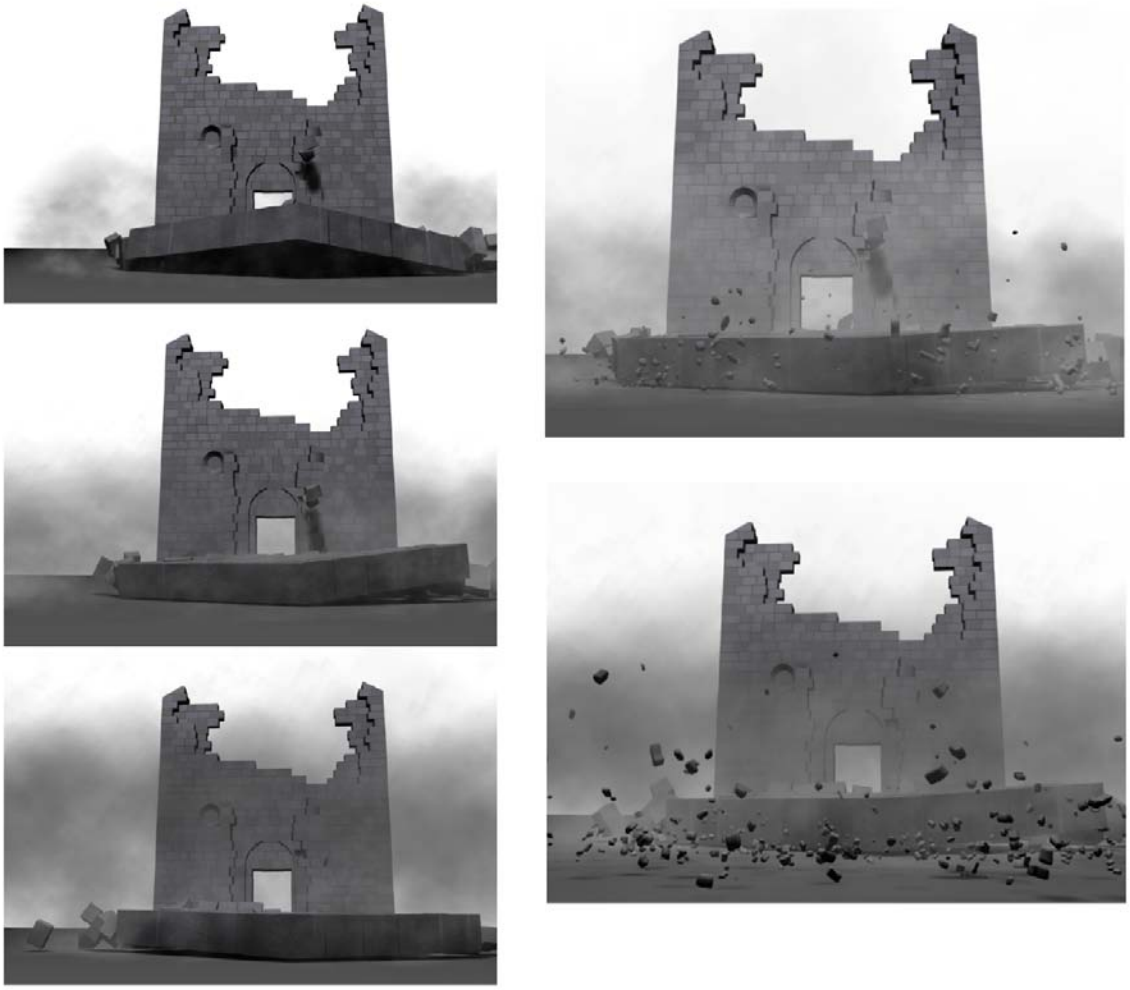


Figure 6: *(left) Some frames from the simulation of the west facade (with textures and dust clouds); (right) some frames from the simulation of the west facade (with textures, dust clouds and fragments)*

tion of infrasound. Therefore, we had to resort to a professional decoder (Lexicon MC-8) that allows for the reproduction of frequencies between 10 and 30 Hz;

- a system for reproducing infrasound. We evaluated existing solutions on the market, from mechanical shakers (i.e., devices able to produce mechanical vibrations, which are usually tied to users' seats and mostly used in theme parks) to more common sub-woofers used in theaters. Mechanical shakers were ruled out due to convenience, complexity and maintenance considerations. We finally considered hi-end hi-fi subwoofers, and chose a VeloDyne DD-18 subwoofer, because it is able to reproduce audio frequencies as low as 10 Hz with the needed acoustic pressure;
- five speakers for medium and high audio frequencies (specifically, TRUTH B2031A).

5. Conclusions

Simulating disasters on cultural heritage objects and building museum experiences about those disasters is an interdisciplinary activity that requires various kind of skills (such as disaster experts, sound technicians, 3D modelers) as well as methods to effectively combine the available sources of information.

In this paper, we have considered the reproduction of earthquakes on cultural heritage buildings, describing the issues we have encountered and the processes we have followed in our project, with the goal of suggesting possible

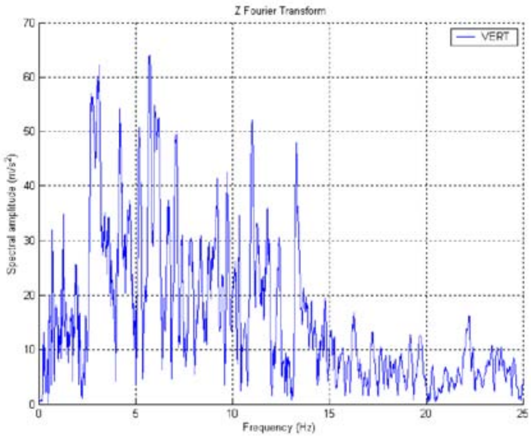


Figure 7: *Graphic plot of the vertical movements produced by the earthquake in the frequency domain.*

methods and solutions for similar projects. While we are aware that simulating completely different kinds of disasters or simulating the effects of earthquakes on much larger scale (e.g., on a city) may require different techniques and approaches, we also believe that this kind of problems are worth being investigated for their different applications (architecture, history, didactic tools for schools and museums, ...).

With respect to future work, we are working on producing a (simpler) version of the simulation that can be rendered in real-time, both as a complement to the museum experience (e.g., to be interacted with after the visitor has seen the movie) and as a multimedia experience that can be downloaded from the internet. Moreover, a 3D stereoscopic version of the movie is being produced, and its introduction into the museum experience will be evaluated in the next months.

6. Acknowledgments

This project has been partially supported by a grant of the "Associazione Sindaci della Ricostruzione del Friuli Terremotato" and the "Associazione Consiglieri della Regione Autonoma Friuli-Venezia Giulia".

Our research has been also partially supported by the Italian Ministry of Education, University and Research (MIUR) under the PRIN 2005 project "VIA - Virtual Interactive Architecture".

Last but not least, we would like to thank a number of people that significantly contributed to the project, in particular Enrico Di Lenarda (3D Modeling Assistant), Silvia Gabrielli (archive footage selection), Alberto Moretti (domain expert), Eric Puntel (simulation assistant), Marcello Riuscetti (domain expert), Augusto Senerchia (sound engineering and editing), and Mattia Vale (storyboarding and 3D animation director).

References

[dHCUCT04] DE HERAS CIECHOMSKI P., ULICNY B., CETRE R., THALMANN D.: A case study of a virtual audience in a reconstruction of an ancient roman odeon in aphrodisias. In *Proceedings of the 5th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage* (2004), Eurographics Association, pp. 9–17.

[GCR01] GAITATZES A., CHRISTOPOULOS D., ROUSSOU M.: Reviving the past: Cultural heritage meets virtual reality. In *Proceedings of the 2th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage* (2001), ACM Press, pp. 103–109.

[MVSL05] MUDGE M., VOUTAZ J.-P., SCHROER C., LUM M.: Reflection transformation imaging and virtual representations of coins from the hospice of the grand st. bernard. In *Proceedings of the 6th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage* (2005), Eurographics Association, pp. 29–40.

[OCA04] O'KEEFE, CIARAN, ANGLISS S.: The subjective effects of infrasound in a live concert setting. In *Proceedings of CIM04: Conference on Interdisciplinary Musicology* (2004), pp. 132–133.

[OH99] O'BRIEN J. F., HODGINS J. K.: Graphical modeling and animation of brittle fracture. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (1999), ACM Press/Addison-Wesley Publishing Co., pp. 137–146.

[OH00] O'BRIEN J. F., HODGINS J. K.: Animating fracture. *Commun. ACM 43*, 7 (2000), 68–75.

[PKA*05] PAULY M., KEISER R., ADAMS B., DUTRé P., GROSS M., GUIBAS L. J.: Meshless animation of fracturing solids. *ACM Trans. Graph. 24*, 3 (2005), 957–964.

[STY*03] STUMPFEL J., TCHOU C., YUN N., MARTINEZ P., HAWKINS T., JONES A., EMERSON B., DEBEVEC P.: Digital reunification of the parthenon and its sculptures. In *Proceedings of the 4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage* (2003), Eurographics Association.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Reflecting on the Creation of an Authentic Aural Experience in the Digital Songlines Game Engine: part of a contextualised cultural heritage knowledge toolkit.

C. Gibbons[1], T. G. Wyeld[2], B. Leavy[3], and J. Hills[4]

[1,2,3,4]Australian CRC for Interaction Design, Australia
[2]IEP, ITEE, Queensland University, Queensland, Australia
[3]Cyberdreaming, Queensland, Australia
[4]Silicon Graphic Incorporated, Queensland, Australia

Abstract

Digital Songlines is an Australasian Cooperative Research Centre for Interaction Design (ACID) project that is developing protocols, methodologies and toolkits to facilitate the collection, education and sharing of indigenous cultural heritage knowledge. The project explores the areas of effective recording, content management and virtual reality delivery capabilities that are culturally sensitive and involve the indigenous custodians, leaders and communities in remote areas of the Australian 'outback'. It investigates how players in a serious gaming sense can experience Indigenous virtual heritage in a high fidelity fashion with culturally appropriate interface tools. This paper describes the circumstances that gave rise to the concept of a 3D ambient audio quilt, designed and implemented specifically for the Digital Songlines game-engine software. It discusses the importance of a site visit to a remote location in the north-east of the Australian outback, and how this prompted the discovery of a new method for creating an authentic aural experience in a 3DVE. This paper reports on completed and ongoing research in this area.

Categories and Subject Descriptors (according to ACM CCS): K.3.1 Computer Uses in Education

1. Introduction

The Australasian Cooperative Research Centre for Interaction Design (ACID) is a collaborative research organisation formed with a number of universities and industry partners. The Virtual Heritage program is a research program under the auspices of the ACID organisation. The digital Songlines project within the Virtual Heritage program is developing protocols, methodologies and toolkits to facilitate the collection, education and sharing of indigenous cultural heritage knowledge across Australian communities, cultural institutions and commercial businesses [GBW06] [LHB*04] [PW06].

The Australian Aboriginal and their culture are known to be some of the oldest in the world. Aboriginal occupation in Australia has been dated at over sixty thousand years, with recent advances and scientific discoveries continuing to change this time frame.

Before 1788 when English settlement commenced in Australia there were approximately 600 languages spoken throughout Australia, with an estimated Indigenous population of 750,000 people [Hen97]. Today Indigenous people make up 2% of the entire Australian population (about 410,000 people). Most of our knowledge of Aboriginal culture is derived from the diverse cultures recorded of relatively modern Aboriginals, particularly those who survived the impact of European colonisation. Hence, the culture of much earlier Australian inhabitants remains problematic [Bic87] [Coo86] [Elk53] [Mem91] [Rid84].

The project objectives are to protect, preserve and promote Australian Indigenous culture, its practices, myths and legends, expanding and re-vitalizing a culture through the visualization of its most prized asset – the land. The project has developed the Digital Songlines software with a virtual landscape encapsulating cultural information, oral histories and mythological stories, based upon the eternal sense of land and spirituality understood by the Aboriginal people of Australia, where feeling, knowing and touching the country, kin and spirit investigating how virtual worlds can capture the spirituality, culture and heritage of Indigenous people and impart these in an empathic way so that non-indigenous people throughout the world can understand the significance and cultural heritage of these areas.

Part of the emphasis on providing a simulated contextually accurate experience of indigenous knowledge is the need for an authentic aural experience within the virtual environment. This paper reports on the experience and outcomes of a site visit to a remote location in the north-east of the Australian outback that prompted a new way of simulating an authentic soundscape.

2. Background – Indigenous cultural heritage

Aboriginal culture was passed onto others through oral traditions, art, dance and rituals. Aboriginal Legends have served an important purpose in the teaching and learning for Aboriginal people, adding to their understanding, connection and interpretation of the world in which they live.

The stories were the means by which knowledge and understanding were passed from generation to generation for over forty thousand years. These ‘yarns’ are vivid, dramatic and informative stories that served the purpose of educating the receiver about all the social, environmental and cultural facts that ensured the ongoing survival and prosperity of the clan. Because they lived with such a close connection to the country and seasons, knew it so intimately, the stories, songs and culture are inextricably linked to the land.

Aboriginal culture is still alive today with older people from the country still able to tell their stories. However, many are passing on and the younger people are becoming lost in the struggle between white and traditional cultures. Some want to know and understand their cultural roots, others want to embrace Western values and deny their heritage. Yet others are simply lost in cultural ambiguity.

In the Digital Songlines project we aim to communicate the culture, history, rituals and stories and association with the country through the 3D virtual worlds, presenting these in the context of the originating country. The importance of this work is in the way it demonstrates an appreciation of the natural environment and the Aboriginal affinity to this land. The virtual world seeks to explore the spiritual, mythic, magic and superstitions of the landscape as a traditional hunting ground and hallowed place of worship.

3. Places of Cultural Significance

To-date, the Digital Songlines project has been used to illustrate significant Aboriginal spaces within Australia including the Mt Moffatt and Carnarvon Gorge National Park areas in south-west Queensland The landscape and surrounding country in these regions is largely undisturbed by modern activities. As such it is a pristine land of gum trees, eucalypts, ironbarks, mulga, caves, granite and sandstone rock formations and fertile, grassy plains. There are innumerable significant places that are marked by a vast array of distinct and special rock art paintings and other cultural artefacts. They contain major meeting places where many different clans descended each year for many months to trade, meet, discuss and follow many practices vital to the survival of the group and the maintenance of their cultural traditions.

4. The ACID toolkit

A core component of the Digital Songlines project is the ongoing development of a digital toolkit. The aim of the toolkit is to be able to effectively communicate meaningful cultural information through a 3D landscape format so the information can be conveyed in context within country. As well as being used for indigenous heritage, this toolkit can also be used to communicate issues of sustainability, land use, water use, explain development issues and contested narrative issues for a number of different uses.

The toolkit needs to be able to facilitate asset management over a large geographical area. This is done preserving high quality local detail. To-date the toolkit features:

1. 3D landscapes based on satellite imagery (with GPS level accuracy at the macro scale);
2. The ability to set weather, time of day or progressive time etc;
3. User level tools to manipulate the landscape and add finer detail at a micro level;
4. The ability to create scenarios and stories and control these through scripts, control of camera position;
5. The ability to create journey paths through the landscape, control the speed and direction along a path;
6. The ability to add flora, fauna related to the area from a database or catalogue of objects;
7. The ability to add ambient audio (wind in trees, bird calls etc), add audio voice over for significant locations (explaining significance of place to viewer, explaining our presence to the spirits etc), and add oral history (automatic or selected avatars);

8. The ability to link to data attributes for presentation of educational material. For example, select information about flora with botanical data; uses such as medicinal; bush tucker; artefact information such as the making of implements for food gathering, or, use as weapons;
9. The ability to participate in massive multi-user serious gaming strategies.

While the primary use of the tool has been in the area of cultural history, a wide range of potential installations have been identified including: museums, science centres, cultural centres, interpretive centres, community consultation, local councils, forestry, water resources, development organisations, schools, mining, safety training, media and data fusion capabilities.

5. Implementation experiences

A highly resolved proof concept prototype has been developed which includes arrays of 3D objects used to recreate a landscape populated by indigenous flora and fauna. These assets have been imported into the game style application based on the Torque Game Engine. The active features include sound, animations, weather and daylight simulation. An established mechanism to import digital terrain models existed and it was modified for importing satellite based geo-spatial data, or data that is prepared for use in GIS software, for accurately mapping the cultural heritage landscape.

Many issues arise from the creation of virtual spaces of some 400 square kms and its reliance on the computational capacity of real-time hardware and visualization technologies. Some are difficult to resolve in a suitable way to communicate the presence required within the virtual space. Such as, how to convey immersive narratologies like: while in place, indigenous knowing pauses at each rock, knows the cycles of the winds, can track underground water, find food and medicine, and uses of the land to speak its stories and keep its history. The kind of knowledge represented and the 'field' in which it is held by local indigenous peoples is often deep, subtle and most intimate [Lan53].

A 'tiered' model has been developed where 'layers' of content are created, accessed, and linked back to the virtual model of the physical place. With such a model, we are able to conceive of the (virtual) land as an interface through which the more traditional dynamics of software creation can be accessed. This layered model allows us to participate in indigenous knowing and being-with, at the most basic level, as the tool is used.

The content can be layered to support virtual heritage applications and narratives (such as land ownership issues, spiritual knowledge, historical and oral stories) and as a community content development and archiving tool (re-populate the virtual spaces with indigenous content). These can be used in entertainment, display, community consultation and education, such as museums, cultural centre displays, as an indigenous language walk, or bush tucker walk, or oral history lesson. These are all developed with the notion of land-as-interface where the (virtual) land is layered with information and practices that arise from that very landscape. In this paper we report on the development and implementation of an authentic aural experience in the digital Songlines game engine.

6. The Limitations of Current Ambient Audio Technology

Ambient audio in most current game engines is represented by either a location based looping soundtrack or by placing static 3D audio emitters around specific nodes of interactivity [Finn04] [Lec02] [Mar01] [McC03] [Wil04]. The design of a looping soundtrack needs to be careful considered so it appears as "dynamic" or randomised sound, and not a loop [Nei05] [NS05] [Roa96] [San04] [Scg94]. For example, when moving through different terrains the user should notice a change in ambient sound levels; a wooded area should sound more alive with wildlife than a sparse terrain. Careful placement of 3D audio objects can significantly enhance the users experience with aural characteristic unique to each area (see figure 1).

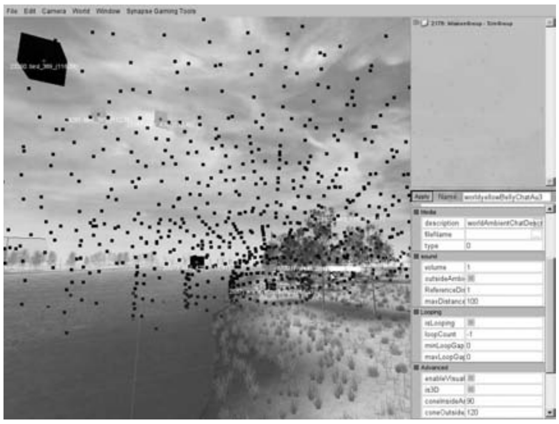


Figure 1: 3D audio emitter in the Digital Songlines Torque Game Engine environment.

7. Collecting Appropriate Audio Assets

By providing the user with multi-sensory awareness information – visual, aural, and tactile (interactive) – a believable landscape simulation experience can be achieved. With the importance of the audio aspect of this virtual landscape experience in mind, the ACID

Indigenous Communities project team embarked upon the collection of a variety of authentic audio ‘assets’ to be used to aurally contextualise a culturally and place-specific 3D virtual environment. A number of locations were identified as suitable. The location reported here is in the remote north-east of Australia. In August of 2005, a field research trip was undertaken to western Carnarvon Gorge in Central Queensland, Australia. The purpose of this trip was to capture the visual and aural environment for incorporation into the Digital Songlines software environment.

7.1 Reflections on the Remote Site Visit: Carnarvon Gorge

As Carnarvon Gorge is a remote area it presented a different aural experience to the urban environment commonly experienced. Most notable was how astoundingly quiet the area is. Such was the extent of this void of sound that quiet sounds, normally obscured through aural masking and filtering, were much more audible. The sound of footsteps on the terrain type being traversed – grasses, leaves, or rocks – could be clearly differentiated with distinct audio differences. These footsteps were capable of dominating the listeners’ audio environment during quiet periods of the day, and could be heard from some distance.

Due to the relative quietness, the acoustic horizon appeared to be much closer than in urban settings. Distant sounds could be heard with greater clarity and definition. For example, the human voice, under certain conditions, could be understood at distances of approximately half a kilometre.

This notion of a closer aural horizon is due to the acoustic properties of the aurally thinner air space in rural environments as there are significantly less audio sources within the listeners’ personal sound field. This reduction in aural density results in distant sounds appearing much closer to the listener, although a perception of distance is still available through the subsequent density of reverberation of the audio source. This raises the question, “how to capture and represent this aural sensation in the Digital Songlines environment?”

7.2 Environment Personality

Prior to the site visit to Carnarvon Gorge the Digital Songlines Engine (DSE) used the more common looping soundtrack to create an aural experience in the virtual environment. However, the very solemnity of the Carnarvon Gorge aural experience prompted us to think about the importance of an authentic aural experience in a new way.

Sitting around the campfire at Carnarvon Gorge, the notion of the environment having a personality emerged. This was triggered by the way the various sounds one could discern in the environment around us at different times of the day seemed to coalesce into what we agreed was an environmental ‘personality’ of sorts. The environment appeared to speak to us through its character. Its character changed as we moved through the environment and at different times of the day. To give this character form we identified specific sounds and isolated them after their recording. In contrast to the quietness of the environment in general, most pronounced was the sheer variety and volume of bird noises. The birds were clearly a large part of the environment’s personality character. At different times of the day particular bird sounds could be heard in the immediate vicinity of our camp. This meant it should be possible (with careful recording and embedment) for our camp environment experience of this personality to be simulated in the DSE 3DVE. The next question was “how to simulate the apparent randomness and dynamism of the changing soundscape?”

7.3 Randomness and Dynamism in a Changing Soundscape

Clearly, as we moved around the camp and into other areas the soundscapes we encountered changed too. The common global, looping, soundtrack would no longer be sufficient. What was needed was a region-specific method for embedding particular sounds that one might encounter in that region, yet dynamically changed to represent different times of the day and to reflect the movement and densities of fauna present.

In its standard format, most game engines do not support algorithmically driven soundtracks. We had to find a method that gave the flexibility and apparent randomness that best reflected the actual environments being simulated. With this in mind the notion of a cellular format was devised.

8. A Dynamic Environmental Soundscape Quilt

By dividing the virtual world space into a series of cells we could recreate the region-specific sounds we encountered from our site visit. These could include sounds embedded in an equally authentic modelled 3D simulation of the environments. The cellular format also allowed specific actions to be performed in relation to an avatar’s movement through the various cells.

The system developed uses a “checkerboard” quilt design methodology with cells monitoring an avatar’s movement throughout the virtual landscape. Upon entering, each cell adjusts the surrounding cells’ audio arrangements – both density and 3D location within constrained random variables. Audio files are randomly selected from a sound bank and used to aurally populate the surrounding cells. The type of audio assets used to populate surrounding cells is dependant on the time of day and any additional required parameters (see figure 2).

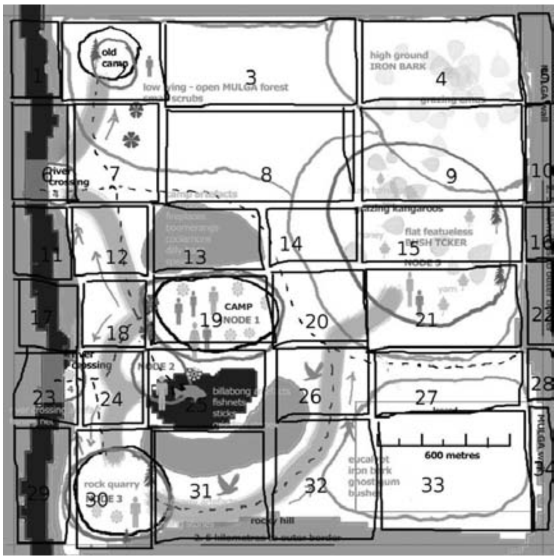


Figure 2: Checkerboard quilt layout of region-specific audio cells.

9. Preliminary Evaluation

The first iteration of this system was evaluated within the project team. Not surprisingly, it proved to be better than what was previously used within the DSE (the looping ambient schema). However, once 3D models and characters were added to the scene, the overhead was too great, resulting in poor performance for both sound and graphical components. Also, difficulty with other animated models within the world, such as birds and non-player characters (such as a human or animal bot in the world that runs on AI and not controlled by a user) through the system triggering these cells. A solution needed to be found that would address these issues.

10. Implementing an Improved 3D Ambient Audio System in the DSE

With the problem of the unacceptably high overhead caused by the first iteration of this system and the triggering action of other objects in the scene, a revised system was developed which leveraged the capacity of the DSE's bitMap code functionality. BitMap codes are usually used for the population of vegetation, specifically grasses, within the DSE world. For example, a bitMap code is used to analyse a prepared .png overlay on the map to determine what type of grass and level of density is needed to populate a specific region.

10.1 Using Graphics Algorithms to Port Audio

We used this semi-random assignment feature of the bitMap code using audio assets in place of models and textures. By implementing the bitMap code for audio, within a controlled radius of the client machine avatar, the ambient audio quilt could be "generated" in real time from a similarly pre-prepared .png overlay. No actual sound generation was taking effect, as every sound was sourced from the prepared sound bank. What was being generated was an algorithmically generated density, placed in a 3-axes coordinated system with a high level of control achieved.

For example, we could place a group of frogs around a water source. We could control the density of sounds, yet the computer dealt with where to place the objects, which sound file to use (from the given sound bank), and randomly constrain their placement within the x y z axes.

The use of the bitMap code functionality reduced the overall overhead and made the DSE navigable again. With this implementation, the ambient audio system can be used for dealing with populations of large and small birds, crickets and frogs on the map, among other collections.

Moving around the map gives one the illusion of different aural soundscapes. When returning to a region, the density and placement of subsequent audio emitters may have changed due to the random nature of the algorithmic system, generating the desired different aural soundscapes.

The system is extendable to handle any additional audio materials, with unique density and placement logic for the algorithms to process and deal with. Combined with looping area effects (such as wind), DSE's Ambient Audio Quilt provides a more accurate aural representation of the landscape than existed under the standard TGE technology.

10.2 A Simulated Aural Experience

With this system we could capture all the sounds of the environment, rendering them algorithmically in real-time. For example, if one walks into a region of bushland in the DSE 3DVE it generates an ambient audio environment in 3D specific to that region of bushland. One might walk through a cluster of trees and hear birds over there, birds over here, but one doesn't know where they are because the sounds are generated algorithmically. What this approach also means is that if one leaves the region and comes back later it will have changed.

This approach came about as a direct result of the experiences encountered on site in the Carnarvon Gorge. It satisfies both our desire for an authentic and accurate simulation of the environment and provides a sensitive setting for the stories to be told.

11. Conclusion

The audio quilt described here provides for an authentic aural experience in the DSE. This forms a critical part of a highly contextualised cultural heritage knowledge toolkit. The importance of contextualising the stories gathered from the community elders is paramount in addressing the sensitivity of their telling. A key tenet of the project is to protect, preserve and promote Australian Indigenous culture, its practices, myths and legends, expanding and re-vitalizing its culture through visualization in a 3D virtual environment. As such, the audio quilt project helps contextualise the virtual landscape with an authentic soundscape where feeling, knowing, touching, and hearing the country, kin and spirit can be experienced.

12. Acknowledgements

This work is supported by ACID (the Australasian CRC for Interaction Design) established and supported under the Cooperative Research Centres Program through the Australian Government's Department of Education, Science and Training.

13. References

[Bic87] Bickford, A. *Aborigines in New South Wales after 1788*. Australia: Australian Government, 1987.

[Coo86] Cook, W. *Aboriginal Involvement in Archaeology*. Australia: Australian Government, 1986.

[Elk53] Elkin, A.P. *The Australian Aborigines*. Australia: Australian Government, 1953.

[Finn04] Finney, K.C., *3D Game Programming All in One*, Premier Press, Boston MA, USA, 2004.

[GBW06] Gard, S., Bucolo, S., and Wyeld, T. G., Capturing Australian Indigenous Perceptions of the Landscape: Virtual environments with cultural meanings, in proc. of, *ALLC'06*, July04-09, Paris, Sorbonne, 2006.

[Hen97] Henderson, J. *Culture and Heritage: Indigenous Languages*. Australia: Australian Government, 1997.

[Lan53] Langloh-Parker, K. *Australian Legendary Tales*. Australia: Angus and Robertson, 1953.

[LHB*04] Leavy, B., Hills, J., Barker, C., Gard, S., and Carrol, J., Digital Songlines: Digitising the Arts, Culture and Heritage Landscape of Aboriginal Australia, in proc. of *VSMM2004*, Ogaki, Japan, Nov 16-22, 2004.

[Lec02] Lecky-Thompson, G. W., *Infinite Game Universe: Level Design, Terrain and Sound*, Charles River Media Inc, Hingam Massachusetts, USA, 2002.

[Mar01] Marks, A., *The Complete Guide to Game Audio: for Composers, Musicians, Sound Designers and Game Developers*, CMP Books, Kansas, USA, 2001.

[McC03] McCuskey M., *Beginning Game Audio Programming*, Premier Press, Boston MA, USA, 2003.

[Mem91] Memmott, P., *Humpy, House and Tin Shed: Aboriginal Settlement History on the Darling River*. Australia: Australian Government, 1991.

[Nei05] Neil, K., Adaptive Audio in Games, Keynote Address and in proc. of *ACMC2005*, QUT, July12-14, 2005.

[NS05] Nordahl, R., Serafin, S., Medialogy and Interactive Sound Design, in proc. of *ACMC2005*, QUT July12-14, 2005.

[PW06] Pumpa, M, and Wyeld, T. G., Database and Narratological Representation of Australian Aboriginal Knowledge as Information Visualisation using a Game Engine, in proc. of *IV06*, London, England Jul05 -07, 2006.

[Rid84] Ridgeway, A. D., *Aboriginal Sites and Involvement of Aborigines in Management and Interpretation*. Australia: Australian Government, 1984.

[Roa96] Roads, C., *The Computer Music Tutorial*, The MIT Press, Cambridge, Massachusetts, London England, 1996.

[San04] Sanger, G. A., *The Fat Man on Game Audio: Tasty Morsels of Sonic Goodness*, New Riders Publishing, USA, 2004.

[Scg94] Schafer, R.M., *The Soundscape: our Sonic Environment and the Tuning of the World*, Destiny Books, USA, 1994.

[Wil04] Wilde, M., *Audio Programming for Interactive Games*, Focal Press, NJ, 2004.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

On the Digital Reconstruction and Interactive Presentation of Heritage Sites through Time

Sabry F. El-Hakim[1], George MacDonald[2], Jean-François Lapointe[1], Lorenzo Gonzo[3], Michael Jemtrud[4]

[1] Visual Information Technology (VIT), National Research Council (NRC), Ottawa, Canada
[2] Bill Reid Foundation, Vancouver, Canada
[3] Centre for the Technological and Scientific Research, ITC-irst, Trento, Italy
[4]School of Architecture, Carleton University, Ottawa, Canada

Abstract

Virtual time travel from existing remains of a heritage site to its previous states and original condition is an educational and interesting experience and can provide better understanding of history. However, digitally reconstructing non-existing objects is a challenge. The interaction and navigation within virtual 4D worlds (adding time to 3D worlds) is also problematical due to the time dimension. In this paper we developed an approach to modelling of heritage sites that has undergone changes over the years. The method creates independent models from different types of data, such as frescos and paintings, drawings, old photos, historic descriptions, and digitization of remains, then assembles and integrates these models for an interactive presentation. Several research issues had to be addressed: (1) Modelling from frescos and drawings with incorrect perspective, (2) modelling from paintings and old photos including fine geometric details from shading (3) colouring models from old photos and drawings to match the colours of existing elements, (4) the seamless and accurate integration of models created independently from different sets of data, and (5) the creation of intuitive interactive presentation that combines all the models and other useful information. We provide contributions to these issues, including our own advanced model viewer, and apply them to modelling of: destroyed Haida house of Chief Weah (Masset, Canada), the demolished and partially relocated Rideau Chapel (Ottawa, Canada), and the Stenico castle (Trentino, Italy) which undergone many changes over several centuries. Each of these diverse examples illustrates different approach for reconstructing heritage sites that changed through time.

1. Introduction

Reconstructing a historical site as it once was or as it evolved over time is one of the most important goals of virtual heritage. The remains of the site, or the site in its present state, either at its original location or where it is currently residing, can be digitized in 3D using digital cameras or laser scanners. Non-existing parts or changed elements have to be reconstructed from any available records such as paintings, old photos, sketches/drawings, written accounts, expert information, and data from similar more complete remains of objects from the same time period. One can anticipate two scenarios:

1- The site/object has changed, damaged, or destroyed.
2- Parts or the entire site have been removed and scattered at various locations such as museums.

Both have modelling and visualisation challenges and require novel presentation to create a virtual site as it has originally been or during other time periods and as it is now, considering that both experts and non-experts are to use the system. The interactivity with models of sites that changed over time adheres to no standard approach and remains a difficult task since there is no simple way to interact with the time dimension [SR02]. Many attempts have been made, however the outcome is usually only a pre-rendered movie, which bypasses the interactivity problems. We address some of the presentation issues, along with the reconstruction of destroyed parts from various records. Specifically, the objectives of this paper are:

1- Modelling of remaining parts and objects at their current location using imaging and laser scanning.
2- Modelling of destroyed parts from any available sources such as old photos, paintings, sketches, plans, and descriptions in historical documents.
3- Assembling and integrating the models to create the site as it originally was and at different periods in its history. Different hypotheses may also be given if there is uncertainty.

4- Designing and implementing interactive presentation with space-time interface to navigate through these models and link with other multimedia components and information related to the history of the site.

Visualising the effect of natural lighting at different day/night time and seasons, and artificial lighting such as candles, is also desirable. The time dimension may not follow a logical linear rule since it is usually compressed and will undoubtedly have missing periods. The remainder of the paper is organised as follows. The next section discusses relevant previous work in modelling and visualising historic sites at various eras, and then elements of our approach are detailed in the third section. Three vastly-different case studies are presented in section 4 followed by concluding remarks.

2. Previous Work

There is a large body of work in each of the issues we focus on. An exhaustive review is beyond the scope of this paper, thus only representative work is given to illustrate the different concepts. We divide the review into projects and modelling and presentation techniques.

2.1. Example of projects

The exhibition Virtual Time Travel in Istanbul [FPM02] comprises the virtual reconstruction of the Hagia Sophia, currently used as a museum. It has been a church and a mosque during different periods in history. This work shows the progression and changes to the architectural and decoration during those periods. The Deep Map system [MZ00] generates personal guided walks for tourists through a city. It takes into account personal interests of the tourist when generating the tour. A 3D reconstruction of the Heidelberg city is part of Deep Map, where destroyed parts can be interactively re-built by the user using a collection of architectural elements (windows, roofs, etc.) that have been used in past centuries. Into the Breath of Sorabol [PKK03] is a virtual event that transformed Kyongju City in Korea into the ancient Silla Kingdom of 1,300 years ago. Interactivity and high-resolution immersive imagery offered seven experiences to the theatre's audiences, including a visit to the Royal Yellow Dragon Temple, 553 AD, and its nine-story wooden pagoda which was destroyed by Genghis Khan and currently only the foundation stones remain. [STY*03] modelled the Parthenon with its carved decorations, most of which are currently in museums, to create a complete virtual reconstruction of the structure in its original and present state. Arrigo VII mausoleum and the partially damaged cathedral in Pisa were reconstructed by scanning the remaining and dispersed pieces [BBC*04]. Lost parts were completed by CAD tools. [SCM04] recreated the now-relocated Kalabasha temple in Egypt, placed it back in its original location and orientation, and illuminated it as it may have appeared some 2000 years ago.

2.2. Modelling techniques from various records

[CRZ00 & ElH01] developed 3D modelling techniques from single images such as old photos and paintings. [SC03] used antique maps to create 3D model and visualise the historical town Kawagoe, Japan. [TYK*04] also used old maps to reconstruct the city of Kyoto over various time periods. [RKN01] used old drawing and experts' accounts to model the historical building Huys Hengelo, Netherlands, which has been demolished and a factory was built on its foundation. [SG00] relied on various descriptions in historical records to reconstruct the Hawara Labyrinth, Egypt, and give the user different possible alternatives of what it may have looked like at 1800 BC. The city of Bologna, Italy, has also been modelled as of today and as it was through different historical periods [Boc04]. Modelling of non-existent buildings was based on historical archives, and archaeological excavations. [BP01] reconstructed da Vinci's ideal city from his sketches and physical models of his machines located at the National Science Museum in Milan. [GRZ04] reconstructed the now destroyed Bamiyan Buddha, Afghanistan, with Photogrammetry using available photos including old high quality ones and tourist photos from the Internet.

2.3. Presentation / rendering techniques

Many of the results of modelling sites at different times are shown in movies where interaction is not possible. Here we focus on approaches that allow user interactivity. There is no standard approach to interact with 4D models, but there are several alternatives that have been proposed, including:

- QuickTime VR object movie [VPW*04]: interactive access to 2D images by rotating the object with horizontal cursor movement and evolving the object through time with vertical cursor movement.
- XML/X3D [HFP06, MGSI01, and Gra02].
- VRML and Java/JavaScript implementation of dynamic virtual environments, particularly with VRML *TimeSensor* node [KHDZ05].
- Combination of techniques such as XML, Java, and panoramas –QTVR [PDH*04].

Approaches to create virtual museums have been summarized by [Bar00 & SLS*05]. They range from Web sites and CD-ROMS to extension of the physical museum with augmented reality or immersive VR. Visualising various types of information attached to the 3D model including uncertainties of the construction hypothesis was studied by [DB03]. Including non-photo realistic or artistic rendering in parts of the presentation to enhance the experience has been advocated [RD03]. 3D game technology has also been suggested [Cha03& LV04]. A VR presentation at the archaeological site of Ename, Belgium [PCKS00] included a 3D model of a destroyed abbey superimposed over a real-time video of the remaining foundation. An accompanying multimedia presentation offered additional information about the site and its past inhabitants. [RKN01] constructed destroyed historical buildings and handled missing parts with conflicting descriptions by giving the user different alternatives using Java and VRML interface that allows switching some parts on and off. [SFR01] developed a set of tools and techniques for visualising and interacting with historical data such as a time lens for viewing arbitrary event times and a time-space exploration tool for simultaneous viewing more than one event time and location of interest. [LS03] used augmented reality technology to overlay real models with 3D virtual models and multimedia contents inside a showcase at a museum. The visitor can interact with the exhibit and see the object throughout its history. [ZCG05] created an interactive graphical user interface with a time slider, along with a time window controlling the start and end of the period, to allow the user to visualise how the site appeared during that period. Through transparency and animation the uncertainty of the construction is integrated in the visualisation. Our approach, presented in 3.3, uses the same idea but adds several other options.

3. Modelling and rendering procedure

The steps we apply for modelling and visualisation of a heritage site through time are as follows (figure 1):

- Collecting material and documents from different time periods and investigating the validity of each (is it a true representation or an artist's conception?)
- Creating 3D models from the documents using modelling from painting and old photos and CAD modelling from drawings and other information.
- Creating 3D models of existing parts with imaging and laser scanning techniques.
- Assembling all 3D models and other data; linking components to each other, correcting scale, filling gaps, and creating smooth transitions.
- Creating an interactive presentation and high quality pre-rendered animations with all models and data.
- Light modelling with different light types at various daytimes and seasons.

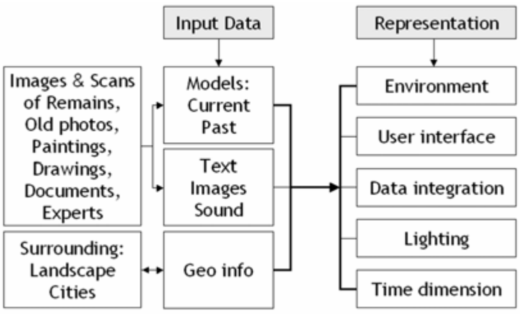


Figure 1: *Data input and representation.*

In the first step, the types of record available vary with each historical period. As far back as medieval time (1200's) one may find surveyors record such as control or markers and their measured locations. In Italy for example those records are called "Libri terminorum". Historic maps are also available for some cities. Pre-Renaissance paintings had incorrect perspective, which means they cannot be directly used for reconstruction. Beginning with the Renaissance, frescos, paintings, and other drawings with correct perspectives were becoming available in addition to accurate line and pictorial maps. In the mid 1800's grey-scale photographs were available and less than a century later colour photographs appeared. Since old visual records cover only portions of the scene, we need to complete the model based on available text descriptions and expert advice and hypothesis using CAD tools. Digital imaging and laser scanning have been widely used, by late the twentieth century, to model existing parts. Details of our approach for 3D modelling are given in [EBP*04]. We will describe some of the procedures next.

3.1. Modelling from medieval frescos/paintings

This procedure is summarised in figure 2.

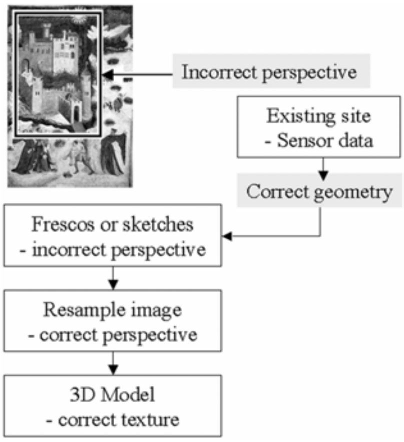


Figure 2: *Modelling from old frescos.*

The method requires that part of the object or site still exists, for example its foundation or archaeological remains. We use measurements from those remains to resample the painting or fresco to create an image with correct perspective and use this to reconstruct the site. If the site has completely disappeared, then there is no reliable procedure and we have to rely on CAD tools to create the model using the fresco mainly as a guide.

3.2. Renaissance paintings and old photos

We use a flexible approach for 3D construction from single images [ElH01]. The approach applies several constraints: coordinate constraints, surface constraints, and topological constraints in two steps: a calibration step and a reconstruction step. The solution is based on the Photogrammetric collinearity equations. Each point p extracted from an image i has two image coordinates, x_p and y_p and contributes two equations:

$$x_p = F_x(f_o, x_o, y_o, X_p, Y_p, Z_p, X_i, Y_i, Z_i, pitch_i, yaw_i, roll_i);$$
$$y_p = F_y(f_o, x_o, y_o, X_p, Y_p, Z_p, X_i, Y_i, Z_i, pitch_i, yaw_i, roll_i) \quad (1)$$

The parameters are three internal camera parameters (focal length fo, and principal point x_o, y_o), six external camera parameters (X_i, Y_i, Z_i, $pitch_i$, yaw_i, $roll_i$), and 3D object coordinates of point p (X_p, Y_p, Z_p). The camera parameters are the same for all points measured in the same image but each point adds three unknown XYZ coordinates. The image coordinates x_p, y_p may also include lens distortion parameters. We solve first for all camera parameters using the constraints: points with same X coordinate, Y coordinate, or Z coordinate, one point with zero coordinates to define the origin of the object coordinate system, and one point with a zero Y and Z to define the orientation. A distance is assigned between two points to define the scale. When sufficient constraint equations are formed together with equations (1), solution of all camera parameters is possible. In the reconstruction step, more constraints are added: shapes such as planes, cylinders, and quadrics, and topological relationships like symmetry and surfaces being parallel or perpendicular. This results in a basic model. A differential shape from shading approach is then used to add fine details to this model. Details of our approach can be found in [ElH06] and it is only summarised her. The process is applied to a work image: a version of the original with some pre-processing such as noise removal filtering and editing of unwanted shades and elements. We first create and triangulate a dense grid of points on the surface of the basic model, then modify the coordinates of these points based on their shading. The amount of modification is obtained from a curve describing the relationship between the grey-levels and the depth difference from the basic model. The curve intersects the grey-level axis at the average intensity value of the points actually falling on the basic model. By adjusting this curve the results can be instantly evaluated and readjusted if necessary. We now have a triangulated grid of points whose coordinates are altered from the initial basic model to account for fine details.

When several images of the same object or site exist, we developed a procedure to create a single consistent model from the images (figure 3). We start with the best image available and create a model as described above. From the resulting model, we now have the geometry that can be passed on to the other images. The model can then be completed from those images in a straight forward manner since all the 3D information has the same scale and coordinate system.

Since old photos are in greyscale, we need to colour them to match the colour of remaining elements. Some techniques are available to transfer the entire colour from a source to the target image by matching luminance and texture information between the images [WAM04]. We adapted a colour-matching tool that matches interactively selected adjacent textures from different images [EGP*03].

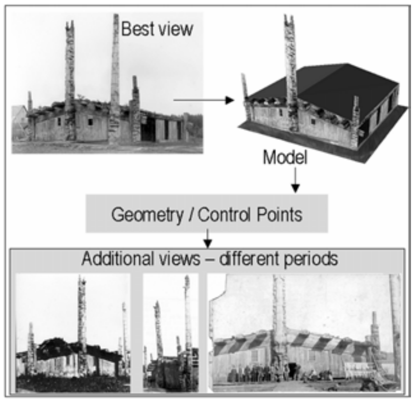


Figure 3: *Consistent modelling from many old photos.*

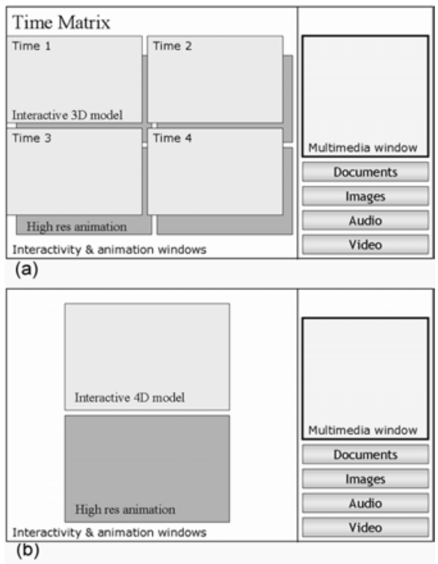


Figure 4: *Two presentation alternatives*

3.3. Presentation / rendering

After consulting our project partners, we opted for two desktop presentation paradigms (figure 4) that are configured according to the project requirements:

- Time Matrix (figure 4a): a different model for each time period that lets the user make the selection from a matrix of displayed windows, or
- A single X3D file that contains all the models from the different times (figure 4b). The manipulation or interaction will be applied to all the models so going from one model to another will be smooth. A time slider is used to select a period in history, and then

opacity or transparency is assigned to model parts based on what should be visible for this time period.

Other information is shown in a separate window and is linked to the model parts in the interactive window where clicking on an object and the desired button invokes the corresponding information. The presentation is prepared in XML with embedded multi-media (video, audio, images, and text). There are two model forms (figure 5): (1) pre-rendered animation with maximum resolution and realistic illumination; and (2) interactive model with resolution to allow real-time interaction. For a large site with more than one building or room, we use a map or floor plan for navigation.

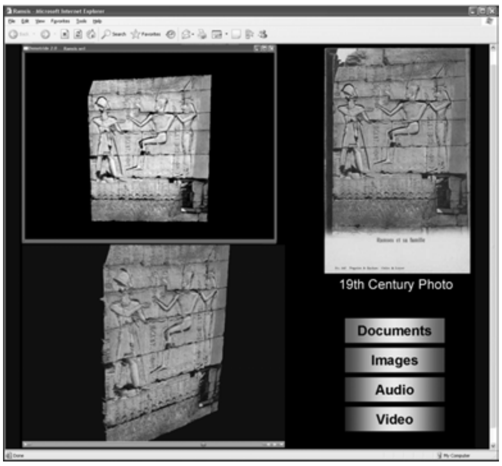


Figure 5: *Screenshot of a presentation.*

3.4. The "Demotride" viewer

To improve user interaction (navigation and object manipulation) with the models, we developed the Demotride viewer [D06], which works with content files based on the VRML and X3D international standards. More specifically, Demotride's support of time sensors allows animation of time and interaction with the 4D models. One feature that distinguishes Demotride from the other viewers is its ability to provide what we call controlled interaction, where the user is bound to respect the visualisation and interaction parameters specified by the content developer. For example, Demotride supports a unique feature that permits specifying the rotational speed of the viewpoint when navigating in the virtual scene. Another interesting feature of Demotride is the support of advanced travel techniques that allow, for example, direct looking at any specific point in the scene located around the user's current viewpoint. These travel techniques developed for virtual walkthrough go well beyond the standard techniques (e.g. rotate, walk, and fly) normally found in other viewers. Other features include support for sound and text and programmable keyboard interaction.

4. Representative Examples

In each of the flowing examples we will indicate the current status of the site, short history, and the available data and records that we made use of to reconstruct it. Modelling of existing parts is detailed in [EBP*04].

4.1. Chief Weah House

The house of Northern-Haida Chief Weah was located in Masset, BC, Canada (figure 6). The structure with its nine skilfully carved totem poles is historically significant and represents a way of life that no longer exists. Known as the "monster" house for its large size, it was built in 1810s and demolished, along with other Haida houses, in the 1880s. A modern house was then built on its foundation. Fortunately part of the foundation still remained in the 1970s under the existing house and some measurements were done then. Also there are written accounts from the 1880's, including some approximate measurements. Many photographs of the house interior and exterior are at the archives of several museums. One of the house totem poles was taken to the Museum of Civilization in Ottawa (figure 7c) where it is now on display. Some articles from inside the house, or similar houses, also survive in museums.

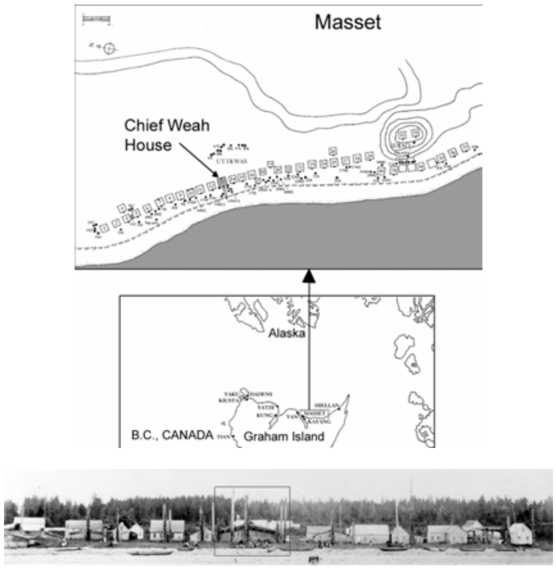


Figure 6: *Masset map and Chief Weah house location (also part of the presentation) and a 1880s photo.*

Totem poles should ideally be scanned to capture all their details. However, due to their size (from 5 to over 10 meters high) and shape, this would be an expensive time consuming exercise for this project. Thus, an image-based method had to be developed [ElH06]. In the first stage of this multi-stage approach, basic shapes are determined with an interactive approach while in the subsequent stages fine details are added automatically by dense stereo matching and shape from shading. The technique also takes advantage of the way totem poles are made: being symmetric and have a flat back. Data from various sources were compiled and used to provide constrains to calibrate the old photos and create the models with the technique described in section 3.2.

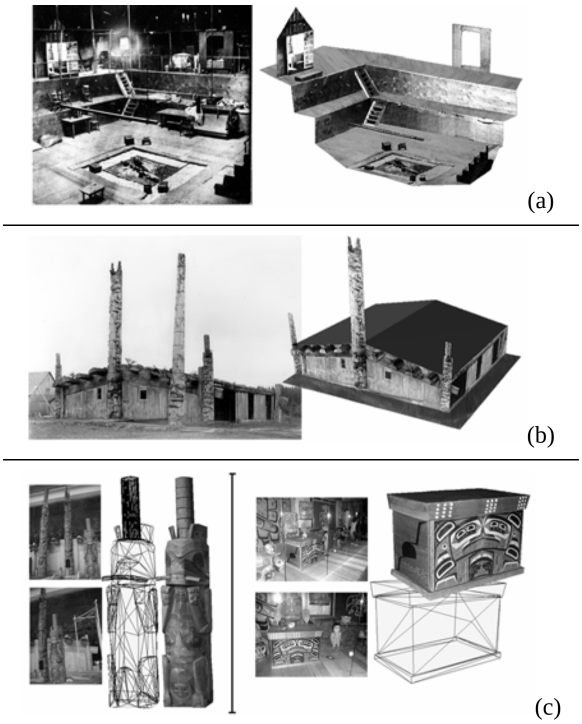


Figure 7: *Haida house models: (a) the inside from old photos, (b) outside, (c) existing totem pole and objects.*

4.2. The Rideau Street Chapel

The chapel of the Convent of Our Lady of the Sacred Heart in Ottawa, better known as Rideau Street Chapel, was demolished by a developer in 1972. Luckily, its architecturally unique interior was dismantled and later reassembled inside the National Gallery of Canada where it is currently preserved (figure 8a). The chapel measured 32 m x 14.5 m x 8.5 m (H). This project objective is to digitise and model the existing interior and reconstruct the destroyed exterior from old images and drawings to create a complete virtual reconstruction of the chapel as it once was. We used: CAD modelling from existing engineering drawings, laser scanning with two different scanners, Photogrammetry, and modelling from old photos. The existing engineering drawings, which were based on surveying and Photogrammetry, created the overall model of the interior of the chapel (figure 8b). More details on the walls and ceiling were obtained by a time-of-flight scanner with 5 mm resolution (figure 8c), and finer details on sculpted surfaces were obtained by a close-range triangulation-based scanner with 0.5 mm resolution. The outside of the chapel was modelled from photographs taken before 1972 (figure 9). All models were integrated together and presented with the tools described in section 3.

Figure 8: *Current chapel interior: (a) overall view, (b) part of wire-frame model, (c) scanned data.*

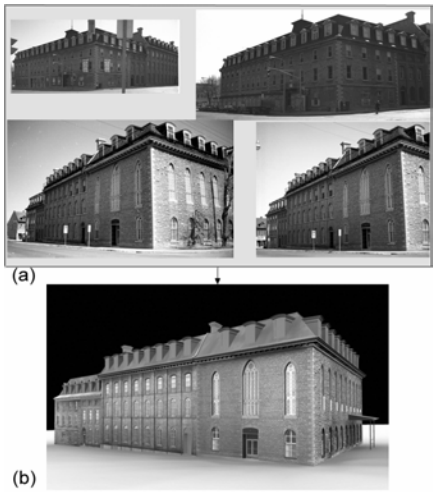


Figure 9: *Rideau Chapel: (a) old images, (b) the model.*

4.3. The Stenico Castle

The Stenico castle is one of the oldest and most important medieval castles in Northern Italy, and is an interesting mixture of styles of buildings added over several centuries. A view of the castle is depicted in the January panel of the “cycle of months” frescoes (figure 10a) in Aquila Tower at Buonconsiglio castle in Trento. The view shows a remarkable difference between the shape of the castle entrance in the 1350s and now. Outer walls and towers have been shortened or removed. We fully modelled the castle as it stands today from aerial and ground images and available floor plans. Using the model of the remains of the altered parts, we reconstructed them as they were from the 1350 fresco using the method outlined in figure 2. The buildings in the current model have also been constructed at different dates, thus the presentation uses transparency on such buildings depending on the time period being observed.

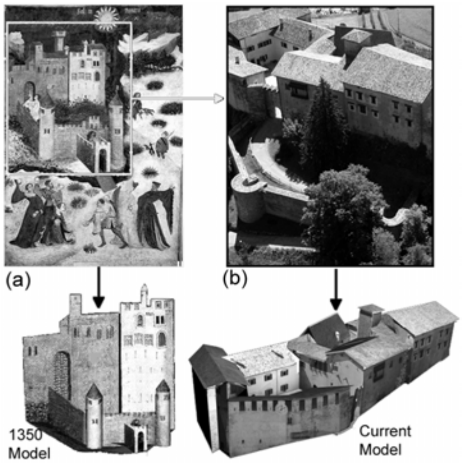


Figure 10: *Stenico Castle models: (a) from fresco, circa 1350, (b) currently from aerial and ground photos.*

5. Conclusions and future work

We have addressed several critical issues in the virtual reconstruction and interactive visualisation of heritage sites that have undergone changes over time or no longer or partially exist. Techniques to create models from frescos, paintings, and old photos were developed and applied to several different projects. Interactive presentation tools to handle 4D models in space and time and other information have been produced. Since we are still refining the procedure, future work includes comprehensive evaluation of the interactive presentation alternatives, at various configurations, by several groups of users. Both experienced and non-experienced users will be drawn upon. Further work involves designing and implementing an immersive system and comparing it to this desktop-based system.

6. Acknowledgements

Charles Hill, art curator of the National Gallery of Canada provided us with unlimited access to the Rideau chapel set up at the gallery. The Stenico castle modelling is funded by the “Provinicia Autonoma di Trento”.

References

[BBC*04] Baracchini, C., Brogi, A., Callieri, M., Capitani, L., Cignoni, P., Fasano, A., Montani, C., Nenci, C., Novello, R. P., Pingi, P., Ponchio, F., Scopigno, R.: Digital reconstruction of the Arrigo VII funerary complex. Virtual Reality, Archaeology and Cultural Heritage, VAST 2004, pp. 145-154.

[BP01] Barbieri, T., Paolini, P.: Reconstructing Leonardo's Ideal City from Handwritten Codexes to Webtalk II: A 3D Collaborative Virtual Environment System. Virtual Reality, Archeology, and Cultural Heritage, VAST2001, pp. 61-66.

[Bar00] Barcelo, J.A.: Virtual museums: When heritage does not exist. World Conference on Science, UNESCO-ICSU and Hungarian Academy of Sciences, 2000, on CD-ROM, ISBN9639046752.

[Boc04] Bocchi, F.: The city in four dimensions: the Nu.M.E. project. *Journal of Digital Information Management*, 2(4), 2004, pp. 161-163.

[Cha03] Champion, E.M.: Applying Game Design Theory to Virtual Heritage Environments. Int. Conf. Computer Graphics and Interactive Techniques in Australasia and S.E. Asia, 2003, pp. 273-274.

[CRZ00] Criminisi, A., Reid, I., Zisserman, A.: Single View Metrology. *Int. Journal of Computer Vision*, 40(2), 2000, pp. 123-148.

[DB03] Dudek I., Blaise J.Y.: Access, interpretation and visualisation of heritage data using the architectural morphology: experimenting emerging interfaces on a case study. ICHIM2003.

[D06] Demotride: http://www.demotride.net, July 2006.

[ElH01] El-Hakim, S. F.: A flexible approach to 3D reconstruction from single images. ACM SIGGRAPH 2001, Technical Sketches, pp. 186-186.

[EGP*03] El-Hakim, S.F., Gonzo, L., Picard, M., Girardi, S., Simoni, A., Paquet, E., Viktor, H., Brenner, C.: Visualisation of highly textured surfaces. Virtual Reality, Archaeology and Cultural Heritage, VAST 2003, pp. 231-240.

[EBP*04] El-Hakim, S.F., Beraldin, J.-A., Picard, M., Godin, G.: Detailed 3D reconstruction of large-scale heritage sites with integrated techniques. *IEEE Computer Graphics and Applications*, 24(3), 2004, pp. 21-29.

[ElH06] El-Hakim, S. F.: A sequential approach to capture fine geometric details from images. ISPRS Commission V Symposium, Dresden, Germany, Sept. 25-27, 2006 (to appear).

[FPM02] Foni, A.E., Papagiannakis, G., and Magnenat-Thalmann, N.: Virtual Hagia Sophia: restitution, virtualization and virtual life simulation. UNESCO World Heritage Congress, 2002.

[Gra02] Grandi, F.: XML Representation and Management of Temporal Information for Web-based Cultural Heritage Applications. *Data Science Journal* 1:1, April 2002, pp. 68-83.

[GRZ04] Gruen, A., Remondino, F., Zhang, L.: Photogrammetric reconstruction of the Great Buddha of Bamiyan. *The Photogrammetric Record*, 19(107), 2004, pp. 177-199.

[HFP06] Hetherington, R., Farrimond, B., Presland, S.: Information rich temporal virtual models using X3D. *Computers & Graphics*, 30(2), 2006, pp. 287-298.

[KHDZ05] Korosec, D., Holobar, A., Divjak, M., Zazula, D.: Building interactive virtual environments for simulated training in medicine using VRML and Java/JavaScript. *Computer Methods and Programs in Biomedicine*, 80 supplement (1), 2005, pp. 561-570.

[LS03] Ledermann, F., Schmalstieg, D.: Presenting an archaeological site in the virtual showcase. Virtual Reality, Archaeology and Cultural Heritage, VAST' 2003, pp. 123-130.

[LV04] Lepouras, G., Vassilakis, C.: Virtual museums for all: employing game technology for edutainment. *Virtual Reality*, 8(2), June 2004, pp. 96-106.

[MZ00] Malaka, R., Zipf, A.: Deep Map - challenging it research in the framework of a tourist information system. 7th International Congress on Tourism and Communications (ENTER 2000), pp. 15-27.

[MGSI01] Mitakos, T., Gergatsoulis, M., Stavrakas, Y., Ioannidis, E.: Representing time-dependent information in multidimensional XML. Information Technology Interfaces, ITI 2001, Croatia, pp. 111-116.

[PKK03] Park, C., Ko, H., Kim, T.: NAVER: Networked and Augmented Virtual Environment aRchitecture; design and implementation of VR framework for Gyeongju VR Theatre. *Computers & Graphics*, 27(2), April 2003, pp. 223-230.

[PDH*04] Pletinckx, D., De Jaegher, L., Helsen, T., Langen, I., Silberman, N., Van der Donckt, M.-C., Stobbe, J.: Telling the Local Story: An Interactive Cultural Presentation System for Community and Regional Settings. VAST 2004, pp. 233-240.

[PCKS00] Pletinckx D., Callebaut D., Killebrew A., Silberman, N.: Virtual-Reality Heritage Presentation at Ename. *IEEE MultiMedia*, 7 (2), 2000, pp. 45-48.

[RKN01] Reidsma, D., Kragtwijk, M., Nijholt, A.: Building "Huys Hengelo" in VRML. ICHIM (2), 2001, pp. 325-332.

[RD03] Roussou, M., Drettakis, G.: Photorealism and Non-Photorealism in Virtual Heritage Representation. Virtual Reality, Archaeology and Cultural Heritage, VAST 2003, pp. 51-60.

[SG00] Shiode, N., Grajetzki, W.: A virtual exploration of the lost labyrinth: developing a reconstructive model of Hawara Labyrinth pyramid complex. CASA Working Paper no.29. Centre for Advanced Spatial Analysis (UCL), London, UK, 2000.

[SR02] Sideris, A., Roussou, M.: Making a new world out of an old one: in search of a common language for archaeological immersive VR representation. Virtual Systems and Multimedia, VSMM 2002, pp.31–42.

[SFR01] Stoev, S., Feurer, M., Ruckaberle, M.: Exploring the past: a toolset for visualisation of historical events in virtual environments. ACM Symposium Virtual Reality Software and Technology, Banff, Canada, 2001, pp. 63-70.

[STY*03] Stumpfel. J., Tchou, C., Yun, N., Martinez, P., Hawkins, T., Jones, A., Emerson B, Debevec, P.: Digital reunification of the Parthenon and its sculptures. Virtual Reality, Archaeology and Cultural Heritage, VAST' 2003, pp. 41-50.

[SCM04] Sundstedt, V., Chalmers, A., Martinez, P.: High Fidelity Reconstruction of the Ancient Egyptian Temple of Kalabsha. Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH 2004), pp. 107-113.

[SC03] Suzuki, S., Chikatsu, H.: Recreating the past city model of historical town Kawagoe from antique map. International Archives of Photogrammetry and Remote Sensing (CD-Rom), Vol.XXXIV-5/W10, ISSN 1682-1777, Vulpera, 2003.

[SLS*05] Sylaiou, S., Liarokapis, F., Sechidis, L., Patias, P., Georgoula, O.: Virtual museums: first results of a survey on methods and tools. CIPA XX Symposium, Torino, Italy, 2005, pp. 1138- 1143.

[TYK*04] Takase, Y., Yano, K., Kawahara, N., Koga, S., Nakaya, T., Kawasumi, T., Isoda, Y., Inoue, M., Kawahara, D., Iwakiri, S.: Reconstruction and Visualisation of "virtual Time-space of Kyoto," a 4D-GIS of The City. Int. Archives of Photogrammetry, Remote Sensing, 35(5), Istanbul, 2004, pp. 609-614.

[VPW*04] Vergauwen, M., Pletinckx, D., Willems, G., Verbiest, F., Van Gool, L., Helsen, T.: As time flies by: Mixed image and model-based rendering of an historical landscape from helicopter images. Virtual Reality, Archaeology and Cultural Heritage, VAST 2004, pp. 241–250.

[WAM04] Welsh, T., Ashikhmin, M., Mueller, K.: Transferring colour to greyscale images. Proceedings of ACM SIGGRAPH, 2002, pp. 277–280.

[ZCG05] Zuk, T., Carpendale, S., Glanzman, W.D.: Visualising Temporal Uncertainty in 3D Virtual Reconstructions. Virtual Reality, Archaeology and Cultural Heritage, VAST 2005, pp. 99-106.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

The backstage of Byblos' Roman theatre: New Digital Devices using Information and Communications Technology (ICT)

G. De Paoli and N. El-Khoury

CAD research group (GRCAO), Faculty of Environmental Design, University of Montreal, Canada

Abstract

This paper presents the results of a multidisciplinary research project that combines the fields of architecture and the conservation of the built heritage, history, communications and computer science. The study of new methods of experimentation will enable us to define and validate new orientations in the way we understand, structure and transfer acquired knowledge about a given architecturally significant complex.

The aim of the project is to present the various experiences obtained during the interpretation of heritage spaces, and in particular intangible heritage, using information and communication technologies. More specifically, it involves acquiring, through ICT, computer modelling and archaeologists' accurate documentation, an understanding of the consequences of successive occupations of an archaeological site on its current condition. It also seeks to gain a better understanding of the construction techniques and know-how of the Ancients.

The objective of this project is to introduce computer modelling, which is capable of showing the site's evolution over the centuries, in order to help us understand the superposition of historic layers.

This work will reflect on how to respond to certain challenges using the example of the experiences acquired at the site of the ancient city of Byblos in Lebanon, a city included in UNESCO's World Heritage List. The Byblos project also helps to re-create and re-mould a monumental complex without having all the information and to test hypotheses that we would otherwise be unable to validate without compromising the heritage values of a site by physically reconstructing it. Such a compromise was experienced in the case of the Roman theatre of Byblos (A.D. 218) which, in the 1930s, was moved and rebuilt by the sea by archaeologist M. Dunand.

1. Introduction

This research project deals with the use of information and communication technologies (ICT) in the enhancement of archaeological heritage, particularly in the case of the Byblos site in Lebanon.

Located north of Beirut, the city of Byblos, which has been included on UNESCO's World Heritage List since 1984, boasts a number of lively, ancient neighbourhoods, as well as an archaeological site where excavation work has unearthed a succession of abandoned cities revealing long-vanished civilizations [Jid04]. "The chief attraction of Byblos for visitors is the superposition, in the same site, of ruins spanning 7,000 years of history." [Dun73]. Over the course of history, the site served as a quarry for successive civilizations. Such was the case with the Roman theatre, which was used as a source of stone by the Crusaders and whose origins date back to 218 A.D. Today, only the first five tiers and the stage remain. When it was first excavated, the theatre faced the setting sun. It was moved and rebuilt near the sea by archaeologists in able to reach the Early Bronze Age period levels.

The idea for this research [DpBEBK05] arose from the study of certain ancient sites, of which sometimes only ruins remain, but that are important enough to merit further exploration: these ruins represent an archaeological heritage. There are multiple definitions of heritage, and this a subject of ongoing debate among the various disciplines involved in its conservation. In our case, we define it as the know-how arising from the construction methods and ways of life that characterize a place through history. The goal is to reconstruct the "memory" of a place through the use of new information and communication technologies (ICT).

The underpinnings of the criteria employed in 1984 in selecting Byblos for inclusion in the World Heritage List have evolved, and consequently the criteria for selecting a site have changed. ICOMOS selected Byblos based on criteria III, IV and VI:

- Criterion III: Byblos offers an exceptional testimony to early Phoenician civilization.

- Criterion IV: Byblos provides one of the first examples of Bronze Age urban organization in the Mediterranean world.

- Criterion VI: Byblos is directly and materially associated with the diffusion of the Phoenician alphabet (on which humanity is still largely dependent), by way of the Ahiram, Yehimilk, Elibaal and Shaphatbaal inscriptions. [IU84]

Heritage is a constantly evolving concept, particularly since 1989 when UNESCO introduced the notion of intangible heritage, the importance of which was confirmed by the "Convention for the Safeguarding of the Intangible Cultural Heritage" of 2003.

The concept of intangible heritage is increasingly gaining currency. According to M. Petzet (2003), "The preoccupation with what we try to define as intangible heritage may also contribute to a broader emotional basis of conservation practice, which can help us in the daily fight against the progressive world-wide destruction and decay of our cultural heritage." [Pet03] Intangible heritage is not a static value, but an evolutionary concept: "The notion of intangible heritage must not be fixed at a particular point in time – it is dynamic and evolves and it is the evolving intangible heritage which is important." [Bum03]

It is also important to note that the concept of heritage was brought to Lebanon in the 19th century by European missionaries, explorers, archaeologists and other scientists interested in identifying and describing monuments from Antiquity and the Middle Ages in the Mediterranean. Since then, the notion of heritage in Lebanon has evolved and now refers to a specifically Lebanese heritage, with ideological undertones directly linked to national identity. Lebanon is composed of several social and denominational identities, which means that heritage is perceived differently according to one's background. Without belabouring these nuances, suffice it to say that heritage and its preservation is a major issue in the world today, and is beginning to gain importance in Lebanon.

These new approaches to tangible and intangible heritage have guided our experimentation in developing new digital devices to protect and promote the archaeological heritage through ICT.

2. Enhancing Heritage Through ICT

One of the main objectives of our research has been to gain an understanding of heritage in order to demonstrate how, using new methods of representation, we can attempt to virtually represent the tangible and the intangible: what is and what used to be. ICTs allow for an exploration involving the simultaneous use of simulation and experimentation, while respecting the guidelines set forth in UNESCO's charters on the safeguarding of heritage.

With regard to techniques of representation, it should be noted that in the twentieth century archaeological heritage was illustrated by means of drawings, plans, cross-sections and elevations produced by architects and archaeologists. Three-dimensional reconstructions of spaces were often created as well. This is precisely what the archaeologist Maurice Dunand undertook during the excavation of Byblos in the 1930s. His sketch of the Roman theatre of Byblos is one example of this approach. It is prominently featured on a sign providing information about the archaeological heritage to visitors (Figure 1). Graphic reconstitution in the form of drawings is still widely used today for the enhancement of archaeological heritage. [DpEk05]

Figure 1: *Sign providing information about the archaeological heritage to visitors (Byblos site).*

The restoration of ancient monuments is another way in which architectural heritage can be enhanced. Restoration, according to Pérouse de Montclos, involves a number of different areas. These include "consolidation", which is done to ensure a building's durability, without modifying it; "reassembly", the reconstruction of a building whose parts are still available on-site; "reconstitution", the collection and reassembly of authentic elements that have been dispersed; and "repair", the replacement of deteriorated parts with new, identical elements. [Gil06]

The physical reconstitution of monuments is now regulated. The Venice Charter, which preceded the one adopted in Victoria Falls in 2003, states "that architectural heritage must be considered within the cultural context to which it belongs, that conservation and restoration of architectural heritage requires a multidisciplinary approach, and that the latter is not an end in itself but a means to an end, which is the building as a whole."

Finally, it should be recalled that, as with life-size reconstructions, scale models also are potential tools for enhancing architectural heritage. They can be used to depict a building, but often provide no information on the factors related to its construction history. Such models are limited to three-dimensional reconstitution of forms that are, in certain cases, hypothetical. They are often used as teaching aids, and also as tools to inform and entertain cultural tourists. [Gil06]

In this chronology of heritage expression, the use of computer models that allow for realistic simulations and real-time movement through virtual spaces, becomes relevant. Little by little, a physical environment is expressed through digital 3D spaces and a kind of architecture that is virtual and dematerialized sees the light and helps us better understand the construction of a building.

3. Methodological Experimentation

In addition to the potential of simulation, virtual space opens new possibilities of expression and experience [And98, [Eng01]. It also helps to re-create and remould a monumental complex without having all the information and to test hypotheses that we would otherwise be unable to validate without compromising the heritage values of a site by physically reconstructing it [UII94]

This was the case of the Roman theatre of Byblos (A.D. 218), which we used as a case. study and which, as previously mentioned, was moved and rebuilt near the sea by archaeologist M. Dunand in the 1930s [Jid04].

In order to better understand the heritage of Byblos' Roman theatre, a preliminary experiment was undertaken based on a multidisciplinary research strategy.

This strategy allows the development of augmented reality experiences in order to make us feel present within old constructions as well as on the actual archaeological site. Based on the 3D models and photos taken from known positions on the site, we built 360° cylindrical panoramic photos related to specific coordinates. Then, through the different configurations of the building, several layers of the theatre were composed on the present site view. The activity of comprehension and communication inside this dynamic virtual historical space is more intuitive and effective than using only abstract technical data for plans and orthogonal views. Furthermore, with ICT, we are able to remotely access these augmented realities. It will also be possible to experience them in an immersive way using projection systems, placing the participant directly in the past while respecting the user scale. In addition, sound effects may be added to enhance the sense of presence. This prototype can also be shared online within the researcher community. [EkDpD06] (Figures 2, 3)

Figure 2: *3D models; corrected perspectives using QTVR.*

Figure 3: *Augmented realities composed with cylindrical panoramic views of the real site; corrected perspectives using QTVR.*

Although many current studies have focused on reading the past using ICT [Ver03, Dav05, GF05, Sal05], the avenues for research identified in the context of this work arose from the activities of the GRCAO, such as the creation of procedural models enabling the modelling of what we have termed "semantic operators". [DpB99]

Such a model allows us to produce figurations with all the characteristics belonging to a single family of objects in order to illustrate the result of a procedure or process, and to find relationships among buildings based on their method of construction; in addition, it makes it possible to verify the validity of a rule characterizing a set of objects, thereby ensuring a dialogue between the various actors and the model. Figuration consists of a series of three-dimensional scenes that act as a metaphor for constructive and temporal reality. [DpB99]

The method of construction can be expressed by means of a functional language that enables the modelling of actions. In the same way, this language can encapsulate and transmit the conceptual properties of an architectural object, and therein lies the strength of the approach that we are presenting. Our reasoning is founded on the generic formalization of a body of knowledge. This descriptive mechanism makes it possible to establish relations between the functions. It helps us to understand the aim of the building's properties through the heritage (memory) of the characteristics that give rise to figurations (Figure 4).

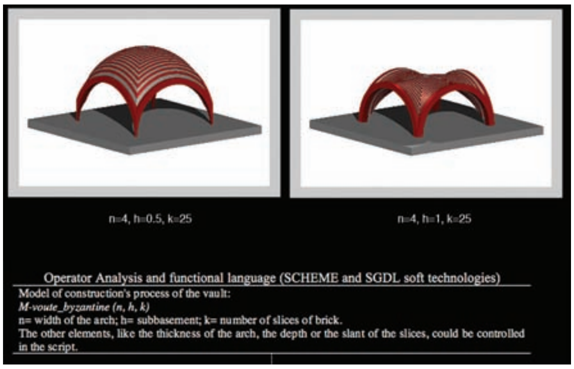


Figure 4: *Example of procedural model.*

Based on these research results, which draw their validity from the know-how of the architectural past, we are able to generate figurations of similar theatres: the researcher operates within a paradigm that brings together various types of solutions by means of "figurations". The use of parametric functions enables us to break down the construction process.

The result is an approach that integrates various design process "operators" to enable the use of appropriately structured figurations that can be manipulated, transformed and organized into a figuration of the object. This process is somewhat similar to the reuse of architectural knowledge in the creation of an architectural work. In designing a new project, architects use fragments of knowledge derived from ancient structures, whether these fragments are reutilized, or simply projected. The design is then based on the memory of a previous interpretation, since the extraction of reusable knowledge can be achieved through an interpretive process. [Leg00]

It is with a view to reusing such knowledge that we take up and further explore this way of representing know-how in the case of the theatre of Byblos.

It is important to note that our intention to transfer ancient know-how does not affect the situation of those researchers who are neither archaeologists nor tourism promoters, but who are attempting to enhance the archaeological heritage. This research project involves several disciplines: architecture, urban planning, history, archaeology, communications and, undoubtedly, other areas as well. The challenge is to bring together all of these competencies around a common objective: the enhancement of heritage. Our goal is to propose methodological experiments to define and validate new orientations in the way we understand, structure and transfer the knowledge associated with a given site.

As a first step, we organized and described approaches for the transfer of knowledge that allow us to exploit representations of spaces with the help of digital modelling software (Figure 5).

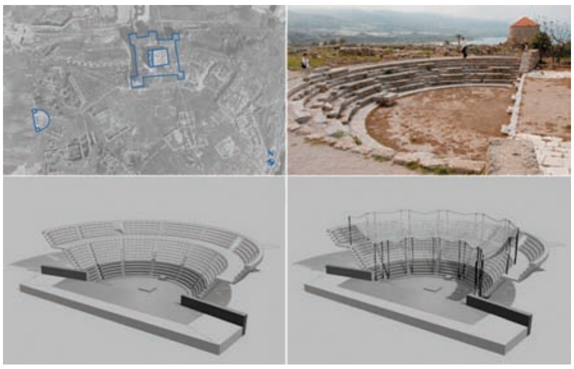


Figure 5: *representations of spaces.*

We then developed the information structure necessary to validate strategies for defining an informative representation model, using software applications that enabled us to create web-based interactive digital devices. (Figure 6)

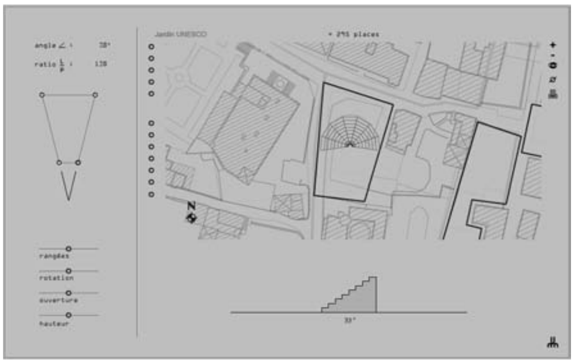


Figure 6: *Web-based interactive digital devices.*

As shown in Figure 7, these experiments enable us to deepen our knowledge of the construction techniques and know-how of the Ancients, and to create computer models that illustrate the site's evolution over the centuries, to help us better understand the superposition of historic layers and suggest periods to which the research proposal might be applied. [DpEk05]

Figure 7: *Site's evolution over the centuries.*

In order to develop modelling methods that take advantage of new information technologies, and to achieve our project's stated objective of designing computer-based approaches for representing the spaces that evoke the memory of a site, we generated an initial computer model of the theatre, in which the steps that have survived to the present (in darker colour) can be distinguished from the rest of the theatre as it might have appeared at the time (Figure 8).

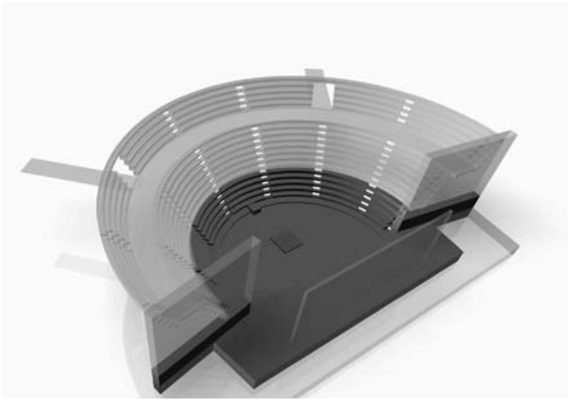


Figure 8: *Computer model of the theatre.*

Models were subsequently created using a 3D digital printer (Figure 9).

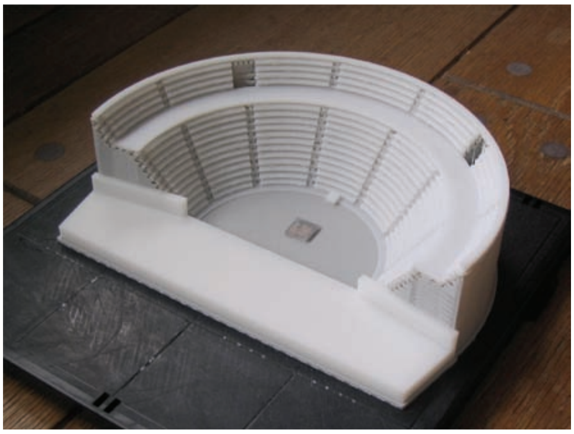


Figure 9: *Model created using 3D digital printer*

Based on this work, the reconstituted theatre was replaced in its original context, as it might have existed before it was taken apart and moved (Figure 10).

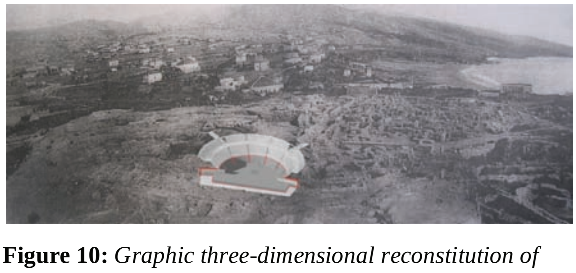


Figure 10: *Graphic three-dimensional reconstitution of ancient site of the theatre.*

The way in which these objectives were achieved will likely have an impact on the current work methods of the designers involved in the development of the Byblos site, as they can now benefit from the integration of information relating to the know-how of archaeologists, architects and historians to colligate a memory and propose new methodological orientations in the restoration of the man-made environment.

4. Conclusion

Through this work we explored different avenues to achieving the goals set forth in the research project. Our initial desire to enhance heritage sites led us to delve further into the concept of heritage and review previous work that had been done on the topic. This reflection process enabled us to define the themes for the Byblos research project by orienting its focus on the case of the Roman theatre. We realized that, although they are two different entities, the archaeological site and the medieval city of Byblos share a common heritage.

We reflected on the types of solutions that could be brought to bear. Our conclusion is that research must not lead to the automatic generation of solutions. Rather, its aim is to provide a means of understanding spaces for the enhancement of local architectural heritage, using information and communication technologies. The reconstitution of a building facilitates not only the understanding of its ruins, but also the dialogue between professionals and non-experts. The creation of 3D models provides a new opportunity for researchers and the public alike to better understand the intangible aspects of heritage spaces.

This work opens up perspectives for research that we will touch on briefly. These can be organized around two distinct poles:

The first relates to the future of heritage education, which is an essential function of a site such as Byblos if it is considered as a museum. From this standpoint, the question of how knowledge will be transmitted in the future is at the forefront, and researchers are asking themselves what type of museum should evolve in the years ahead. Will we see an "info-highway museum" in which data can be manipulated but not modified? Or a "museum without walls" that will be able to reach people wherever they are? [AB97]

Heritage education is a key element in the creation of an open museum space, as in the case of Byblos, which opens onto the city. Knowledge is thus transferred in a kind of museum entertainment.

The second pole is the adoption of an "interdisciplinary vision in the development of conservation and outreach initiatives to better reflect the richness and subtlety of the content." [AB97] The various disciplines involved in the enhancement of heritage —history, archaeology, architecture, urban planning and others—come together and generate debate and discussion to propose solutions that integrate their combined expertise.

Research perspectives are not limited to these two poles but extend to other areas as well. For instance, the results of this research project could be used for pedagogical purposes. Teaching the history of vanished heritage sites could be made more accessible by means of models to aid in our understanding of these sites.

It is in this direction that we continue our methodological experimentation; our aim is to contribute to the advancement of knowledge by integrating the expertise of the various disciplines involved in the enhancement of heritage sites, which can only be enriched as a result. ICTs become a unifying element between these different fields and thereby help to facilitate the process of shared decision-making.

References

[AB97] ALLARD M., BOUCHER S.: Prolégomènes au développement de modèles théoriques de pédagogie muséale. In *Le musée, un lieu éducatif,* Musée d'art contemporain de Montréal, 1997.

[And98] ANDERS P.: *Envisioning Cyberspace : Designing 3D Electronics Spaces*, McGraw-Hill, 1998.

[Bum03] BUMBARU D.: Patrimoine matériel et immatériel, devoir et plaisir de mémoire. In *La mémoire des lieux : préserver le sens et les valeurs immatérielles des monuments et sites,* Actes du Symposium scientifique international, ICOMOS, Victoria Falls, 2003 `http://www.international.icomos.org/victoriafalls2003/kimberley_fre.pdf`

[Dav05] DAVE B.: Labyrinthine Digital Histories, Interpretive, Extensible and Referential. In *CAAD Futures 2005, Learning from the Past*, Springer, 2005, 53-62.

[DpB99] DE PAOLI G., BOGDAN M.: The Front of the Stage of Vitruvius' Roman Theatre - A new Approach of Computer Aided Design that Transforms Geometric Operators to Semantic Operators. In *CAAD Futures 1999, Computers in building*, Kiuwer Academic Publishers, 2005, 321-333.

[DpBEBK05] DE PAOLI G., BRONSON S., EATON L., BONNECHÈRE P., KHAYAT G.: *Expériences méthodologiques et validation de nouvelles orientations basées sur des dispositifs numériques pour la création d'un espace-mémoire pour l'ancienne ville de Byblos au Liban.* (2005), Research project CRSH-IDR Canada, 2005.

[DpEk05] DE PAOLI G, EL KHOURY N.: ICT and the Ancient City of Byblos : A new Direction for the Communication of Intangible Heritage. In *Augmented Heritage,* Europia Production, 2005.

[DpEk05] DE PAOLI G., EL-KHOURY N.: L'arrière-scène du théâtre romain: le cas de Byblos au Liban. In *Virtual Retrospect 2005,* Collection «Archéovision» de l'Institut Ausonius, 2005, 18-22.

[Dun73] DUNAND M.: *Fouilles de Byblos,* (5 volumes), Librairie orientaliste Paul Geuthner – Librairie d'Amérique et d'Orient, Adrien Maisonneuve, 1939-1973.

[EkDpD06] EL-KHOURY N., DE PAOLI G., DORTA T.: Digital Reconstruction as a means of understanding a building's history: case studies of a multilayer prototype. In *Communicating Space(s),* (Proc. eCAADe2006), 2006

[Eng01] ENGELI M.: *Bits and Spaces : architecture and computing for physical, virtual, hybrid realms*, Birkhäuser, 2001.

[GF05] GUTIERREZ D., FRICHER B.: Virtual Crowds in a Digital Colosseum. In *Virtual Retrospect 2005*, Collection «Archéovision» de l'Institut Ausonius, 2005, 82-87.

[Gil06] GILLOT C.: La restitution du patrimoine archéologique bâti, aux moyens des technologies de l'information et de la communication, État de l'art In *Rapport de recherche GRCAO, Contribution d'un moyen informatique à la validation d'hypothèses de restitution architecturale dans le domaine du patrimoine bâti ancien*, University of Montreal, CRSH Canada, 2006

[IU84] ICOMOS UNESCO: Liste du patrimoine mondial N. 295, Bien proposé: Byblos, Liban, ICOMOS, Paris, 1984 `http://whc.unesco.org/pg.cfm?cid=31&id_site=295`

[Jid04] JIDEJIAN N.: *Byblos à travers les âges*, Librairie orientale, 2004.

[Leg00] LÉGLISE M.: Conception assistée : modélisation et interprétation. In *Modélisation architecturale et outils informatiques entre cultures de la représentation et du savoir-faire*, Les Cahiers scientifiques N. 95, Acfas, 2000. 51-66.

[Pet03] PETZET M.: La mémoire des lieux:préserver le sens et les valeurs immatérielles des monuments et sites. In *Place, memory, meaning: preserving intangible values in monuments and sites,* Proceedings 2003 `http://www.international.icomos.org/victoriafalls2003/papers.htm`

[Sal05] SALEH F..: CULTNAT & The art of Documenting Heritage. In *Virtual Retrospect 2005,* Collection «Archéovision» de l'Institut Ausonius, 2005, 106-112.

[UII94] UNESCO, ICCROM, ICOMOS: Document de la Conférence de Nara sur l'Authenticité. In *Convention du Patrimoine Mondial*, Nara, 1994.

[Ver03] VERGNIEUX R.: Virtual Reality: a Tool for Archaeologis. In *Virtual Retrospect 2003,* Collection «Archéovision» de l'Institut Ausonius, 2003.

The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

A Parametric Exploration of the Lighting Method of the Hagia Sophia Dome

W. Jabi[1] and I. Potamianos[2]

[1]New Jersey Institute of Technology, USA
[2]Democritus University of Thrace, Greece

Abstract

Byzantine church design depended heavily on natural light for the generation of evocative effects supportive of the liturgical acts taking place in it. It appears that there were several effects coordinated in order to strengthen the impression of divine presence within the church. The method of ritual lighting reveals a sophisticated level of design in Byzantine churches involving a number of issues which must have formed a complete and integrated system. Specifically, one of these effects concerns the generation of light shafts within the church apse and a derivative apse geometry; a second effect concerns the lighting of the dome; a third effect concerns the church proportions in respect to the way its interior spaces ought to be viewed as well as a number of other relevant issues. Some of this work has been presented in various papers, conferences, and speeches. However, recently we have employed parametric modeling as a tool that helps us comprehend more fully and accurately the design strategies and methods involved as well as revise erroneous assumptions made in the initial stages of this research. The present paper aims at exploring the effect of the luminous dome through the aid of a 3D model focusing specifically in the system developed for the initial dome of Hagia Sophia in Constantinople. While this idea has been previously explored theoretically, the 3D model provides us with more accurate results, verifies some of our conclusions and refutes others becoming in this way not only an indispensable tool that allows us to reach a more detailed and clear exploration of our initial assumptions.

Categories and Subject Descriptors (according to ACM CCS): I.3.5 Computational Geometry and Object Modeling and I.3.7 Three-Dimensional Graphics and Realism

1. Introduction

Byzantine church design depended heavily on natural light for the generation of evocative effects supportive of the liturgical acts taking place in it. It appears that there were several effects coordinated in order to strengthen the impression of divine presence within the church. The method of ritual lighting was quite complicated and reveals the quite sophisticated level of design in Byzantine churches involving a number of issues which must have formed a complete and integrated system. Specifically, one of these effects concerns the generation of light shafts within the church apse [POT1996] and the dependence of the apse geometry on these [PJ2006]; a second effect concerns the lighting of the dome [POT2004]; a third effect concerns the church proportions in respect to the way its interior spaces ought to be viewed [PJT1995] as well as a number of other relevant issues [POT1996]. Some of this work has been presented in various papers, conferences, and speeches. However, recently we have resorted to the aid of parametric modeling which provides us with a tool which helps us comprehend more fully and accurately the design strategies and methods involved as well as revise erroneous assumptions made in the initial stages of this research. The present paper aims at exploring the effect of the luminous dome through the aid of a 3D model focusing specifically in the system developed for the initial dome of Hagia Sophia. While this idea has been explored theoretically before [POT2004] the 3D model provides us with more accurate results, verifies some of our conclusions and refutes others becoming in this way not only an indispensable tool but also pressing us to achieve a more detailed and clear exploration of our initial assumptions. For instance, the accuracy achieved in a recent exploration of the light shafts generated in the church apse was impressive. The parametric model we created helped us discover the heptagonal geometry of the Hagia Sophia apse in plan and the reasons for its adoption, an issue that has never been undertaken before by architectural historians [PJ2006]

The focus of this paper concerns the manner in which the central dome was lit in domed church types. Even today, Byzantine church domes often appear constantly radiant especially around the apex area. The radiance of the dome is enhanced by the shape of the pendentives that are also lit and appear to dynamically lift the dome off the ground (Figure 1). It imparts the impression of a luminous complex launched from four points at the base and stretching over the interior space as if it were a weightless membrane not subjected to the law of earthly gravity, a notion carrying clear symbolic connotations. The image of Pantocrator, often placed at the apex, is not lit directly by sunlight but, since he is considered to be the distributor of "true light" to the world, is made to appear to emit constant light. Practically, such an impression could not occur by chance. Instead it would be the result of a specific study by Byzantine architects.

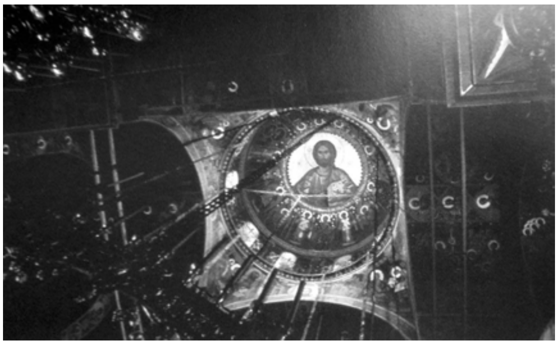


Figure 1: *Iveron Monastery Church dome and pendentives, Mount Athos, Greece.*

A study of this kind seems to have been undertaken in Hagia Sophia resulting in an impressive solution. Written evidence to this is provided in the extant writings authored by its architect Anthemius of Tralles in the form of solid geometry propositions for the construction of reflectors.

One of these reflectors was designed for the purpose of directing solar light at a single point within a building, taking into account the varying positions of the sun throughout the day and the seasons (Note 1) (Fig. 2).

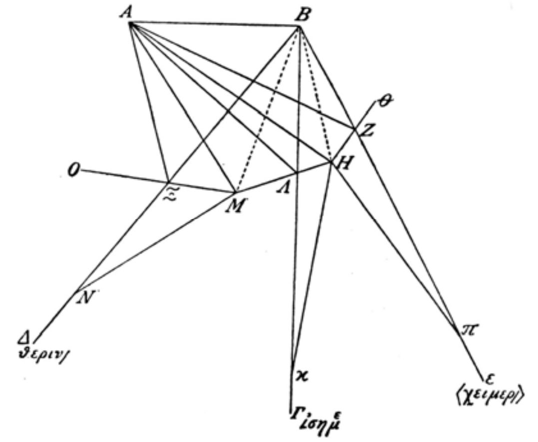


Figure 2: *Anthemius' reflector $\Theta ZH\Lambda MN\Xi O$ reflects sunlight coming from either Δ, Γ, or ε, and passing through B, to A. Δ represents summer solstice, ε winter solstice and Γ the equinox.*

As it is revealed by these written sources, Anthemius was an expert in the handling of light, being able to direct, focus, and stabilize it. It should be expected, thereby, to use his capacities if necessary in designing Hagia Sophia, his life masterpiece. The manner of the admission of light, as it appears, was critical to the new church, a fact which not only led Anthemius to creating reflectors of local importance but, as we shall see, played a significant role in the shaping of major formal elements as well. In fact, the handling of light became a decisive generative force of the geometry and an integral part of the built form.

Regarding this particular reflector, there are clear indications both in the drawing and the accompanying text that it was intended for Hagia Sophia. These indications have been thoroughly analyzed elsewhere (Note 2).

Anthemius states: "It is required to cause a ray of the sun to fall in a given position, without moving away, at any hour or season." (Note 3)

In order to satisfy the conditions posed, the reflector would have to be elliptical in three dimensions. It would have to possess, that is, a double curvature. Any mathematician could ascertain that the form of an elliptical reflector of this kind would be independent of latitude and as long as light penetrated through the center (B) of the prescribed opening, i.e. the one focus of the ellipse, would end up to the second focus (A). Although the reflective properties of the ellipse were known to Anthemius (Note 4) he resorted to a new solution which he related to the solstices and the equinox. Instead of using the simplest and most direct method for constructing an ellipse by means of a string attached to its two foci (A & B), Anthemius resorted to a quite complicated process of resolution by means of tangents, which was used here for the first time (Note 5). The reason was, as it seems, that this resolution presented certain practical advantages (Note 6). The "practicality" in which the problem was solved was strongly criticized by modern historians of science as evidence of Anthemius' limited mathematical expertise (Note 7).

The usefulness of the ellipsoidal reflector becomes clear when one reads the only description of Hagia Sophia available from that era, contained in the work of Procopius, Justinian's court historian (Note 8). This account provides evidence that light did not penetrate uncontrollably into the space but entered through deep windows lending an "incessant gleam" to the dome. This is precisely what would have happened in case the ellipsoidal reflectors had been placed at the window sills. The apex of the dome would have been constantly and brilliantly illuminated. Furthermore, if these three dimensional reflectors did not occupy only the window sills but continued up on the window reveals on both sides of the window, then the effect of a brilliant apex would have remained constant throughout the day commencing at sunrise and lasting until the sunset. In order for this effect to be achieved, the window sills and reveals would have to have acquired a certain depth. In this way no direct light would have been able to enter but only indirect or reflected. As far as its construction is concerned this reflector would consist of flat mirrors subdivided into ever smaller parts until they reached the nearest possible approximation to an ellipse, that is, by golden, glazed mosaic laid out on a preset mortar base. This method was frequently used at the time and had reached a level of unprecedented refinement (Note 9).

But since by using these reflectors virtually the apex of any dome could be brightly and constantly lit no explanation is provided regarding the form of the original dome itself which, as reported by contemporary historians, was of an unusual and structurally vulnerable shape. All contemporary accounts agree that it had a precariously flat curvature which, as modern scholars have repeatedly speculated, might have caused its subsequent downfall. At least one of them asserts that this was due to the absence of sound practices of structural engineering [WAR1976]. It would seem reasonable, however, that the architect's choice should be attributed not to his lack of knowledge, in which case he would have followed time-honored antecedents, but, instead, to his will to defy technological limitations in order to achieve a highly desired effect.

While the original dome no longer survives, we have historical descriptions of it and thus certain theories have been proposed regarding its form. Although these vary somewhat regarding its plan, they agree on its vertical profile (Note 10). The impact of the slight plan variations would not have been considerable in respect to the effect of light. On the contrary, the shallow profile, of which we are certain, would have had a decisive impact on the way light would be captured and reflected by its surface (Note 11).

The explanation regarding the vertical profile of the dome is probably found in another reflector design by Anthemius (Note 12). The significance of this reflector design lies on its similarity to the original dome profile, on the one hand, and to a peculiarity observed in the process of its solution, on the other. This problem deals with the issue of the reflection of light on the concave side of a spherical catoptrical surface (Figure 3).

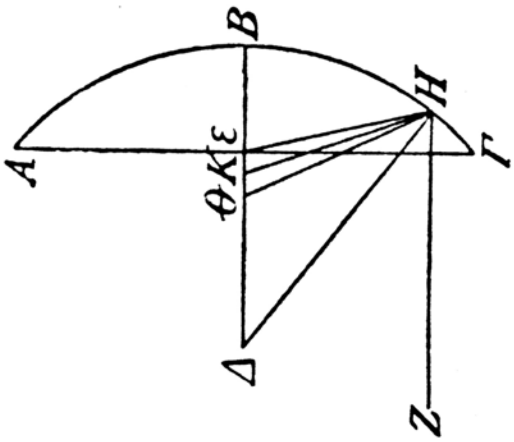


Figure 3: *Anthemius' reflector ΑΒΗΓ reflects any ray parallel to ΔΒ, such as ΖΗ, to Κ.*

The reflector ΑΒΗΓ is given the shape of the circular segment circumscribed around a square of side ΑΓ. In this circular segment the solar rays that fall with a direction perpendicular to the chord ΑΓ will be reflected toward point K.

Discussing some preliminary issues of historical interest about the problem at hand, Anthemius mentions that Apollonius had already discovered the focal point but that he himself had some additional suppositions to prove. While Anthemius examines the focusing capabilities of the spherical reflector in relation to those of the parabolic, which he has dealt with in the preceding problem he offers no reason for choosing this particular circular segment. Obviously, the focal point would remain the same regardless of the size of the circular segment. But if the segment were larger, that is, if it were a semicircle then due to its greater size the focusing strength of the reflector would be doubled. But no such thing is mentioned.

The peculiarity noticed in the problem solution is the following. After Anthemius has proved that the focal point is K, curiously enough, he returns to dealing with the subject in order to prove the equality of the angle that the incident ray creates with the curvature ZHΓ to the corresponding angle between the reflected ray and the curvature KHB. This proof does not add anything to the preceding one but it clarifies the equality of the angles of incidence and reflection the rays form with the curvature *per se* rather than with the tangent. In this manner he interrelates the motion of light rays with the curvature without any intermediaries such as the centre, the radius or the tangent in order to ensure that whatever the angle of incidence, the corresponding angle of reflection will be governed by a simple and constant rule. There is only a short distance to the observation that the more obliquely light comes in the larger the number of successive reflections will occur within the circular segment (Figure 4). This probably constitutes the additional hypothesis that Anthemius is referring to, but unfortunately the text stops before the final conclusions are presented.

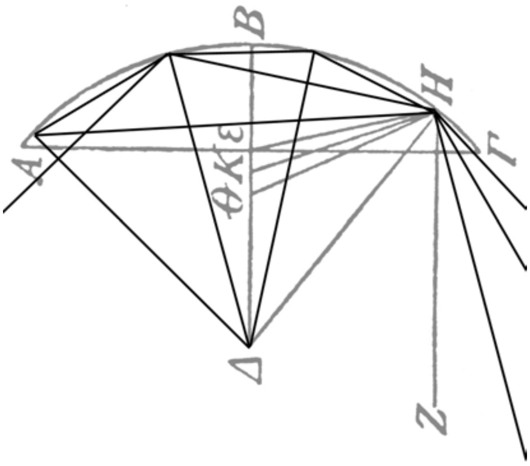


Figure 4: *An illustration of light behavior when incident rays are oblique (superimposed over Anthemius' original reflector diagram).*

By giving the dome such a shape, it would be possible to receive intensely oblique light rays and create multiple reflections within the dome lighting both the southern and the northern areas of it. In achieving this oblique penetration of light the role of the ellipsoidal reflector would be decisive. For instance, the rays that would present 15º-30º divergence from the perpendicular to the chord would generate a second reflection within the curvature of the reflector while with a divergence greater than 30º would generate a third reflection. If one considered that a large number of rays would be introduced with an even greater obliqueness, that is, around 45º, this complex of reflectors would cause an entrapment of light which would generate a uniform illumination of the dome.

The curvature profile of the original dome (Note 13) in transverse section is identical with the arc shape described by Anthemius' problem (Note 14). It is especially remarkable that the base of the inscribed square coincides exactly with the lower cornice while the upper horizontal line of the square coincides precisely with the cornice at the base of the dome windows (Figure 5).

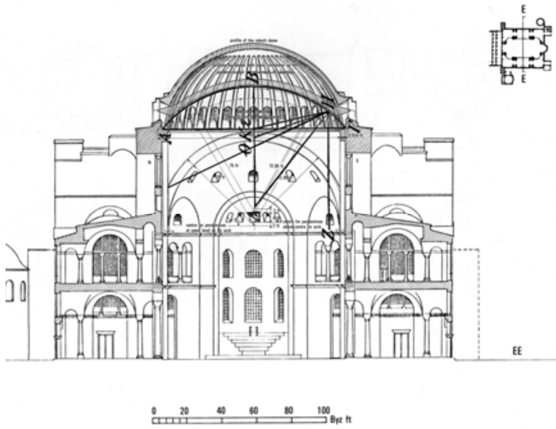


Figure 5: *Anthemius' reflector diagram superimposed over a cross-section of Hagia Sophia (with original and current domes). (Source: Anthemius and Mainstone).*

By accurately placing Anthemius' proposition for a spherical reflector over the two overlaid drawings of the initial dome and the present one by Mainstone, we discovered some additional relationships. First, that Δ by coinciding with one of the possible centers of the initial dome (since there are several theories about its exact height) indicates the most probable dome profile. Second, that HΘ, where Θ is the middle of ΔB determines the limiting case for the reflection of rays coming parallel to ΔB and at the same time determines the lowest point of the pendentive. This means that the light reflected from the church floor as the most extreme case would be reflected first onto the dome and from there no lower than the springing of the pendentive, a principle which would ensure a gradual increase of light intensity from the pendentives toward the dome. Finally, no reflections from the floor would exceed Hε , i.e. the level of the cornice, therefore there would be no disturbance of the illuminated dome by any motion taking place on the church floor.

The words of Procopius might be indicative of this entrapment of light when saying that it: "… is singularly full of light and sunshine; you would declare that the place is not lighted by the sun from without, but the rays are produced within itself, such an abundance of light is poured into this church." (Note 15)

Thus the effect of the combination of this dome profile with the ellipsoidal reflector would ensure that the dome would receive light from the greatest possible number of windows at any one time thus providing a constant and brilliant light at the apex throughout the day and the year while the shallow dome curvature would generate light reflections that would be trapped within it driving away all shadows and achieving an even illumination. The apex would be enveloped in a sort of a cloud of light. The specular surfaces of golden, glazed mosaic surrounding all window reveals and window sills would create a bleeding effect on the window sides so that the distance between the windows would be perceptually diminished disconnecting, in this way, the dome from its base.

Thus, the cupola would appear completely detached from the rest of the church, creating the impression of an area of light existing on its own accord. This light would be perceived as an unnatural, unearthly light, entirely detached both physically and metaphorically from the present world being isolated, suspended and hovering above, unconnected to the events and occurrences of the world below. It would be a wonderful representation and a most convincing impression – in spatial terms – of that indescribable divine light, the "superessential" light, the light "that casts no shadows".

2. The construction of the parametric model and the lighting model

In order to verify the hypotheses put forward above, we made use of sophisticated parametric software to reconstruct and verify the geometric proofs and relationships put forward by Anthemius. Additionally, to test the light behavior, we reconstructed the basic geometry of the original dome and used an accurate lighting model to observe its behavior.

2.1 The parametric model

In order to verify that Anthemius' reflector (as shown in Figure 2 above) actually works, his textual instructions were followed to create a parametric model using Bentley's Generative Components (GC) software. The GC software allows the user to define global parameters (variables) and establish complex geometrical relationships among entities in a 3D environment. We used GC to construct the mechanics of Anthemius' reflector with the ability to modify the angles and the distances at will. The model behaved as expected and does verify that light passing through point B will indeed be reflected to point A (Figure 6). The next step was to superimpose the parametric model over the cross-section of the Hagia Sophia dome by Mainstone to observe if the reflector geometry is consistent with the architectural environment. It is important to note here that because the original dome has collapsed the exact geometry of the dome as well as the parapet and cornice underneath it is unknown (Figure 7).

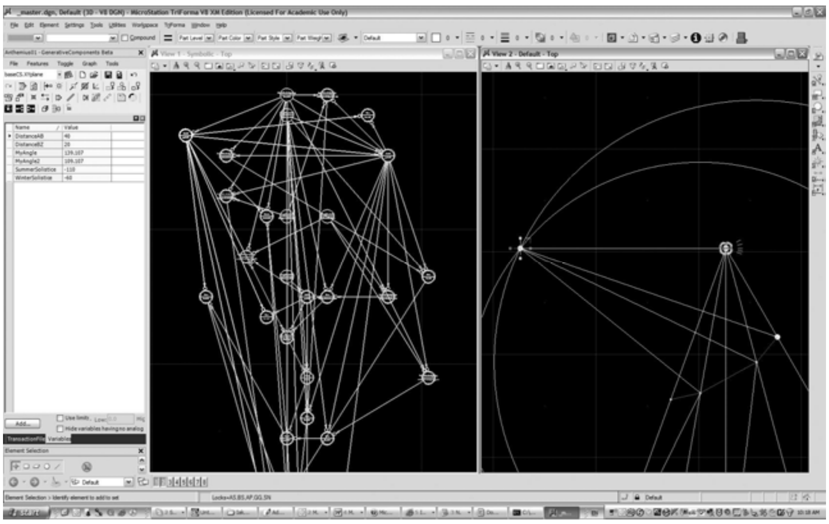


Figure 6: *Screen capture of the parametric model reconstructing Anthemius' reflector.*

Thus, the exact parapet geometry as shown in Mainstone is to a large extent conjectural as is the one we are proposing. However, as the reader will notice in the next section, Mainstone's parapet proposal creates an uneven lighting condition that is not consistent with the descriptions of the dome from that period.

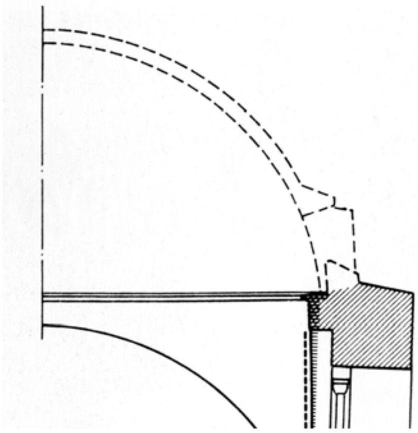


Figure 7: *Illustration showing that no parapet remained after the collapse of the dome. (source: Mainstone)*

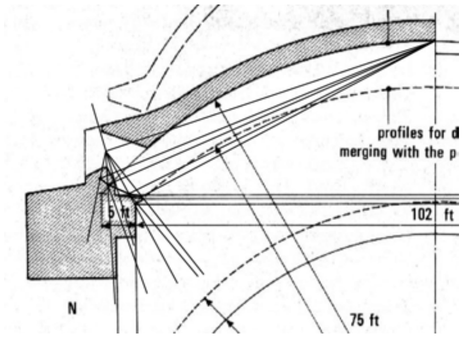


Figure 8: *Parametric model of Anthemius' reflector superimposed over the cross-section of the original dome (source: Mainstone).*

2.2 The lighting model

To test the hypothesis that the flatter geometry of the original dome combined with the geometry of an ellipsoidal mirror at the base of the dome created a more uniformly lit dome with a continuous ring, we modeled Mainstone's proposal for the original dome. Computer-based lighting models have been shown to be accurate tools [GTG1984]. We used AutoDesk's 3D Studio MAX and applied the most accurate advanced lighting method (Radiosity combined with ray-tracing with high-accuracy settings). As an initial test, we constructed a simple room with a window and rendered it with and without a mirrored surface (Figure 9). We used this model to ensure that the recipient surfaces are indeed affected by the simple addition of a reflective surface into the scene. We then used the same materials and applied them to the dome geometry. We also used the same lighting model (Accurate IES Sky and IES Sun light). Given the limitations of time and scope of this paper, we did not vary the hour and date. We are planning a more complete investigation in the future.

Next we constructed the geometry of the dome and two versions of the parapet (the one proposed by Mainstone) and the one proposed by us as illustrated in Figure 8 above (Figure 10). We then rendered each, viewing the dome from the bottom looking up at the dome (Figure 11). Using Adobe Photoshop, we applied the same curve distortion to both original images to better discern the light distribution in the dome (Figure 12). It is clear that the ellipsoidal model created a brighter and more uniform light distribution. The geometry of the curved mirror and the flat geometry of the dome combined to create an effect of an "incessant gleam" as suggested by accounts from that period. The solution proposed by Mainstone provided a much more uneven lighting condition with artifacts due to the sudden change from an inclined cornice to a vertical parapet.

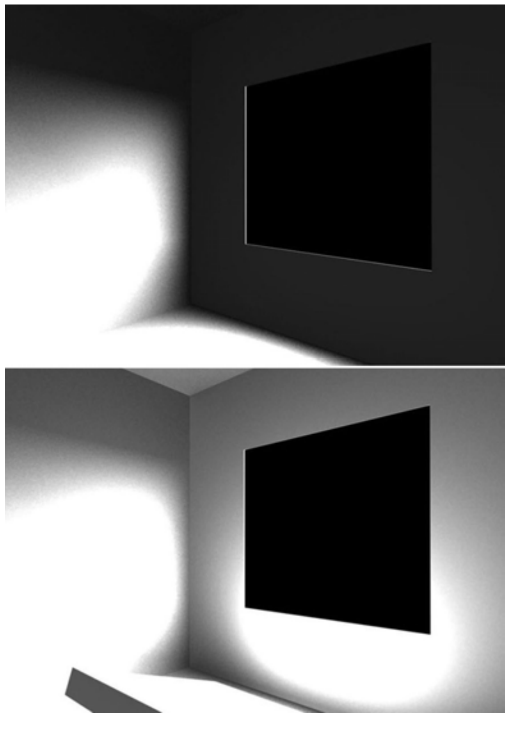


Figure 9: *Lighting model test.*

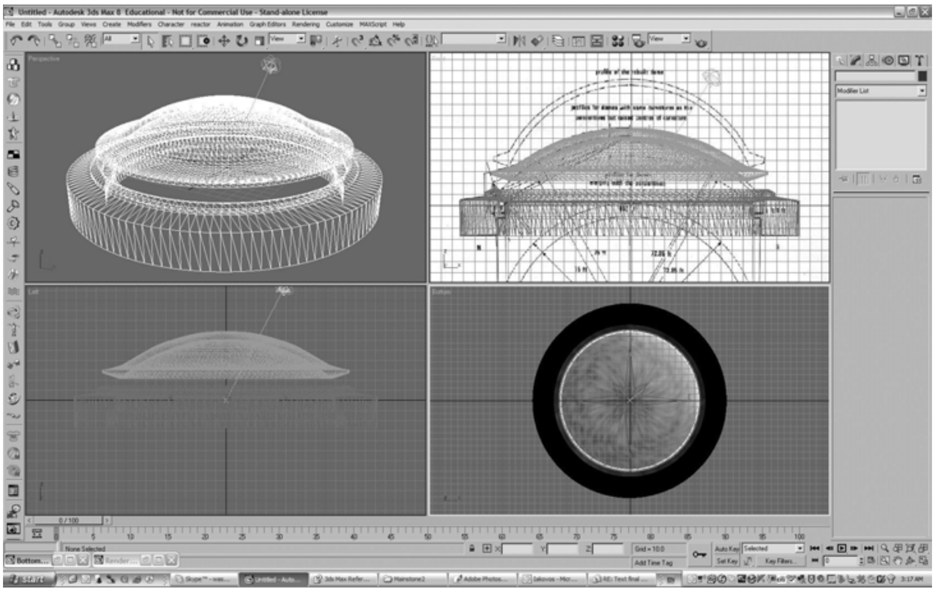


Figure 10: *Screen capture of the 3D Studio MAX model of the dome.*

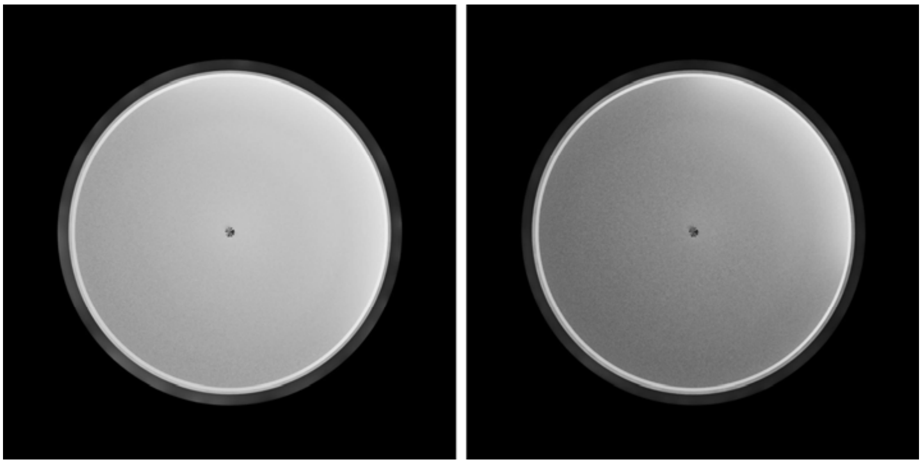


Figure 11: *Lit domes with ellipsoidal mirror (left) and slanted cornice (right).*

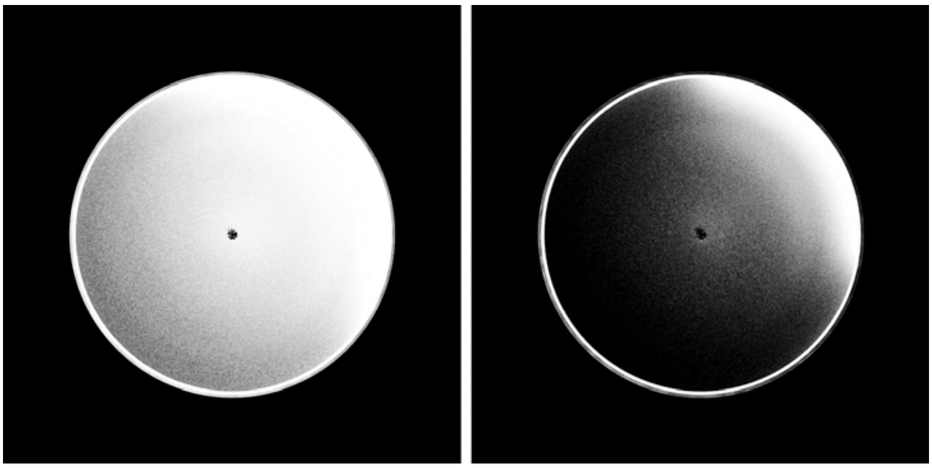


Figure 12: *Curve distortion applied to lit domes with ellipsoidal mirror (left) and slanted cornice (right).*

3. Conclusion and future work

Light and geometry played an important role in the design of Byzantine churches and in the design of Hagia Sophia in particular. The advent of sophisticated parametric digital tools and advanced lighting models has allowed us to test many of the theoretical propositions and discover new findings regarding the behavior of architectural constructions. The model allowed us to synthesize Anthemius' research on reflectors with the original geometry of the window sills and the dome of Hagia Sophia. Furthermore, it also provided an accurate visualization of the lighting conditions within the dome given two possible profiles for the parapet and cornice. This visualization could then be compared to historical documentation and other evidence.

Many of the results presented in this paper require further verification and elaboration. Future work will include the design of more sophisticated 3D parametric models and the testing of other churches and noted works of architecture. The dome of Hagia Sophia requires further elaboration of the pendetives to fully account for their role and lighting. The geometry of the windows at the base of the model is largely unknown and requires further investigation to determine its attributes and resulting effect on the lighting of the dome.

The parametric tools are of value in of themselves for analysis as well as design. We intend to develop generalized parametric tools that will help the designer specify desired conditions and explore alternatives created through the manipulation of parameters and geometric relationships. Such digital tools allow us to think algorithmically and parametrically about our designs such that we can create spaces and solids that are constructed rigorously and exploit the full potential of geometry and light.

Notes

1) The geometrical solution to this problem appears in his excerpt entitled Peri Paradoxon Michanimaton (About Paradoxical Machines) reproduced and with an English translation in [HUX1959], pp. 6-9.

2) See [POT1996] and [POT2004].

3) See [HUX1959], p. 6.

4) See [HUX1959], p.10.

5) See [HUX1959], pp. 9-10.

6) These advantages are analyzed in [POT1996] and [POT2004].

7) See [CAM1990], p. 121 and [TOO1976], pp.187-20

8) See [PRO1961], I, I, 41-3. For a full analysis see [POT2004].

9) See [MAT1964], p. 29.

10) See [SWI1940], p. 153.

11) See [MAI1988], p. 127.

12) This design is contained in the excerpt Fragmentum Mathematicum Bobiense, also attributed to Anthemius in [HUX1959], pp. 22-3.

13) See [CON1939], p. 589 and [MIC1976], p. 37.

14) See [HUX1959], p. 22.

15) See [PRO1896], p. 6.

References

[CAM1990] CAMERON, A.: Isidore of Miletus and Hypatia: On the Editing of Mathematical Texts. Greek Roman and Byzantine Studies 31 (1990), 103-127.

[CON1939] CONANT, K.: The First Dome of St. Sophia and Its Rebuilding. *American Journal of Archaeology 43* (1939), 589ff.

[CON1946] CONANT, K.: The First Dome of St. Sophia and Its Rebuilding. *Bulletin of the Byzantine Institute* (1946), 71-78.

[DOW1960] DOWNEY, G.: *Constantinople in the Age of Justinian*. University of Oklahoma Press, Norman, OK, 1960.

[FOS1852] FOSSATI, G.: *Aya Sophia*. London, 1852.

[GTG1984] GORAL, C., TORRANCE, K., GREENBERG, D., BATTAILE, B.: Modelling the interaction of light between diffuse surfaces. In *Proc. 11th annual conference on computer graphics and interactive techniques*, ACM Press (1984), 213-222.

[GRU1958] GRUMEL, V.: *Traité d'Études Byzantines, I: La Chronologie*, Bibliothèque Byzantine, Presses Universitaires de France, Paris (1958),163-165.

[HEA1921] HEATH, T. L. Sir.: *A History of Greek Mathematics*. Oxford: Clarendon Press (1921). Greek translation, 2001).

[HOL2002] HOLME, A.: *Geometry: Our Cultural Heritage*, Springer-Verlag, New York, 2002.

[HUX1959] HUXLEY, G. L.: *Anthemius of Tralles: A Study of Later Greek Geometry*. Greek Roman and Byzantine Monographs, Eaton Press Inc., Cambridge, MA, 1959.

[KAL1946] KALLIGAS, M.: *He Aisthetike tou Chorou tes Ellenikes Ekklesias sto Mesaiona*, Athina.

[MAI1988] MAINSTONE, R. J.: *Hagia Sophia: Architecture, Structure and Liturgy of Justinian's Great Church*. Thames and Hudson, London, 1988.

[MAI2001] MAINSTONE, R. J.: *Hagia Sophia: Architecture, Structure and Liturgy of Justinian's Great Church*. Thames and Hudson, New York, 2001.

[MAT1964] MATHEW, G.: *Byzantine Aesthetics*. John Murray, London, 1964.

[MAT1971] MATHEWS, T. F.: *The Byzantine Churches of Istanbul: A Photographic Survey*. Pennsylvania State University Press, 1971.

[MIC1976] MICHELIS, P. A.: *Hagia Sophia*. Athina, 1976.

[POL1999] POLYBIOU, M.: To Katholiko tes Mones Xeropotamou, *Ypourgeio Politismou* (1999), Athens, 19.

[POT1996] POTAMIANOS, I.: Light into Architecture: The Evocative Use of Natural Light as Related to Liturgy in Byzantine Churches, *Dissertation* (1996), University of Michigan, Ann Arbor.

[POT2004] POTAMIANOS, I.: The Mathematics of the Ideal Dome. In, *Aesthetics and Architectural Composition - Proceedings of the Dresden International Symposium of Architecture*, Weber R., and M. Amann (eds.), (2004), pp. 66-71.

[PJ2006] POTAMIANOS, I., JABI, W.: Interactive Parametric Design and the Role of Light in Byzantine Churches. In *Communicating Spaces, Proceedings of the eCAADe conference* (2006), University of Thessaly, Volos, Greece, September 6-9, 2006. (Forthcoming).

[PJT1995] POTAMIANOS, I., JABI, W., TURNER J.: Exploring the Proportions of Middle Byzantine Churches: A Parametric Approach. In *Proc. Sixth International Conference on Computer-Aided Architectural Design Futures* (1995), pp. 483-493.

[PRO1896] PROCOPIUS: *Of the Buildings of Justinian*. Palestine Pilgrims' Text Society Library, London, 1896.

[PRO1961] PROCOPIUS: *De Aedificiis*. Loeb Classical Library, Harvard University Press, 1961.

[SWI1940] SWIFT, E. H.: *Hagia Sophia*. Columbia University Press, 1940.

[TOO1976] TOOMER, G. J.: *Diocles on Burning Mirrors: The Arabic Translation of the Lost Greek Original*. Springer-Verlag, 1976.

[TRO1928] TROWBRIDGE, M. L.: *Philological Studies in Ancient Glass*. University of Illinois Studies in Language and Literature, 13, (3-4), 1928.

[VAN1965] VAN NICE, R. L.: *Saint Sophia in Istanbul; An Architectural Survey*. Dumbarton Oaks Center for Byzantine Studies, 1965.

[WAR1976] WARREN, J.: Greek Mathematics and the Architects to Justinian. *Art and Archaeology Research Papers*, Coach Publishing, London (Dec.), 1976.