National Research Council Canada | Conseil national de recherches Canada

Institute for Information Technology | Institut de technologie de l'information

*Exploring a byzantine crypt through a high-resolution texture mapped 3D model: combining range data and photogrammetry. **

J-A Beraldin, M. Picard, S.F. El-Hakim, G. Godin, C. Latouche,
V. Valzano, A. Bandiera

September 2002

* published in: Proceedings of the Intern. Workshop on Scanning for Cultural Heritage Recording - Complementing or Replacing Photogrammetry, Corfu, Greece. September 01-02, 2002. NRC 44916.

EXPLORING A BYZANTINE CRYPT THROUGH A HIGH-RESOLUTION TEXTURE MAPPED 3D MODEL: COMBINING RANGE DATA AND PHOTOGRAMMETRY

J-A Beraldin [a, *], M. Picard [a], S.F. El-Hakim [a], G. Godin [a], C. Latouche [a**], V. Valzano [b], A. Bandiera [b]

[a] Institute for Information Technology, National Research Council Canada, Ottawa, Ontario, Canada, K1A OR6
[b] SIBA Coordination, University of Lecce, Lecce, Italy, 73100 - siba@siba2.unile.it

Commission V, Working Group V/2

KEY WORDS: Cultural Heritage, Virtual Reality, Laser scanning, Three-dimensional, Reconstruction, Texture, Registration, Photo-realism.

ABSTRACT:

In recent years, high-resolution recording of heritage sites has stimulated a lot of research in fields like photogrammetry, computer vision, and computer graphics. Numerous algorithms and methodologies have been proposed in the literature. In practice, what a 3D photographer needs is a commercially available solution to this so-called as-built documentation. In this paper, we present an effective approach for photo-realistic 3D model building from the combination of photogrammetry and 3D range data. The method is applied to the virtualization of a Byzantine Crypt where geometrically correct texture mapping is essential to render the environment accurately in order to produce enticing virtual visits, apply virtual restoration techniques on the frescoes and remove architectural elements that have been added over the years so that the site can then be viewed in the correct historical context. A movie entitled "CARPINIANA" was created in order to demonstrate the results.

1. INTRODUCTION

The capacity to create, display, manipulate, archive and share a digital representation of the shape and appearance of an existing object (as-built documentation) finds a most challenging class of applications in high-resolution recording of heritage-related objects and sites. Once the object or site has been "virtualized", one can use immersive technologies to study or to promote a cultural site using a virtual 3D visit. However, beyond photo-realistic rendering, a 3D model contains a wealth of information that can be analyzed and enhanced. Features that are small or only visible from a distance can be interactively examined, thus, allowing the study of fine details such as tool marks or surface texture. Furthermore, sites that must be closed for conservation reasons can still be studied and visited once a 3D model has been created. Computer-based visual enhancement and analysis techniques can be applied to the digital model in all of these situations. One such application is found in virtual restoration of an historical site. As opposed to "traditional" restoration that is performed on the physical object or site (usually not reversible), virtual restoration is applied directly onto the digital copy hence reversible. For instance, it allows the optimization of the legibility of textual and artistic informative data, without turning to interventions often traumatic for the original copy (SIBA Web Site). Or in some instances, architectural elements that have been added over the years can be removed and the digital 3D model of a site can then be viewed in the correct historical context.

As a way to demonstrate the proposed modeling method, we selected a Byzantine Crypt (see Fig.1). Though not part of a typical tourist itinerary in Italy, this rupestrian site contains amongst the oldest Byzantine frescoes that signed and dated. It is known as the Crypt of Santa Cristina and is located in Carpignano Salentino in Apulia (Bandiera 1980). This Crypt measures about 16.5 m by 10 m by 2.5 m. This one thousand year old crypt presented many challenges from the technical and historical point of view. The Crypt also contains an altar and three pillars added during the Baroque period. In order to model a complete site like this Crypt, a 3D photographer would have to be skilled in a number of 3D modeling procedures. The 3D photographer could be a specialist that does this type of work on contract basis. Or, in cases where the property of data is of concern, it can even be a technician directly link to the agency requesting the work. Whichever case it may be, the combination of range data with photogrammetry was examined from a user point of view. Numerous papers have dealt with this issue, but one problem remains to be addressed i.e. the availability of commercial tools for both 3D modeling and high-resolution texture mapping onto dense 3D models. Very few solutions exist on the market addressing this issue. We report an effective approach based on commercial software to the problem of high-resolution photo-realistic texture mapping onto a 3D model generated from range images. Section 2 presents an overview of the processing pipeline and section 3 describes the elements used in that pipeline. Section 4 shows how these elements were put together for the Byzantine Crypt. Finally, concluding remarks appear in section 5.

a) b)

Figure 1. Byzantine Crypt *IX-X*, a) the two outside entrances, b) view of the interior located underground.

* Corresponding author. J-A Beraldin, angelo.beraldin@nrc.ca, http://www.vit.iit.nrc.ca NRC 44916
** Currently completing its Master's thesis (Université Laval, Qué, Canada) as a guest worker at the NRC Ottawa

2. GENERAL OVERVIEW OF THE PROCESSING PIPELINE

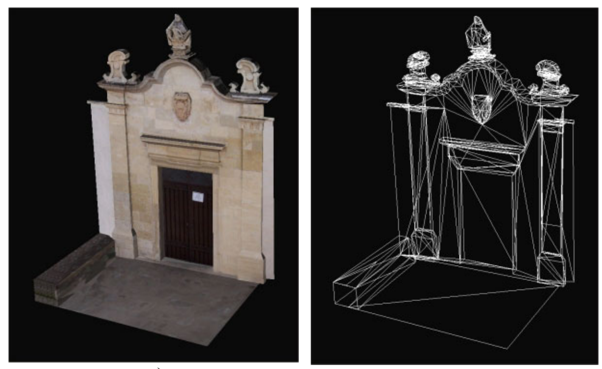
The typical processing pipeline used for 3D modeling includes calibration/verification, geometric modeling, and appearance modeling. The sequence of steps required is well documented in (Soucy et al., 1996a). Here we summarize some of the steps for the reconstruction of a complete fully textured model. The *calibration* of a range camera is concerned with the extraction of the internal parameters of the 3D camera. The manufacturer should include with their commercial 3D camera a test object to verify the accuracy (verification). For 2D cameras, different calibration methods exist spanning from the simple pinhole model to the complete photogrammetric solution.

Geometric modeling is essential to recreate realistic models of sites that have a lot of surface details difficult to model with photogrammetric techniques. In most cases, the creation of a 3D model will rely on multiple scans (range images), taken at various locations all around an object or inside a site, that need to be registered. A few techniques have been devised for this problem. One method combines photogrammetry and laser range-imaging techniques (El-Hakim et al., 1998). Another method uses only the surface data from the multiple views (Soucy and Laurendeau, 1995). The views must have enough overlap between them to find the registration and to merge them together. One important feature of this approach is that the geometric details of the surface of the object itself are used to register the views together. Obviously, quasi-planar or spherical surfaces should be avoided with this latter technique and for those cases, the former method, or the following one, is recommended. Some 3D camera manufacturers have adopted the use of geometrical objects like spheres placed near a surface to be acquired in order to facilitate the registration between the individual 3D images.

Appearance modeling includes methods like image perspective techniques (IPT) and reflectance modeling. IPT is concerned with direct mapping of photographs onto a 3D model (Weinhaus and Devarjan, 1997; El-Hakim et al., 1998; Neugebauer and Klein, 1999; Sequiera et al., 1999; Stamos and Allen, 2000). Reflectance modeling is used to extract from the measured colour and shape those physical properties of an object that are intrinsic to it and that determine its appearance when viewed with artificial lighting on a computer screen (Baribeau et al., 1992; Bernardini et al., 2001). Texture mapping is also an efficient way to achieve realism with only a low resolution, faster to render, geometric model. Recently, techniques that map real-scene images onto the geometric model, also known as image perspective techniques (IPT) have gained a lot of interest. Though, some commercial 3D systems supply a colour texture in registration with the 3D image unfortunately with very limited visual image quality. Hence separate cameras acquire high-resolution colour images, which can be precisely mapped onto the geometric model provided that the camera position and orientation are known in the coordinate system of the geometric model. The main challenges faced by people in that field are adequate lighting, accurately computing lens distortions, 2D camera to 3D-model pose estimate, dealing with hidden surfaces and incomplete views (El-Hakim et al., 1998).

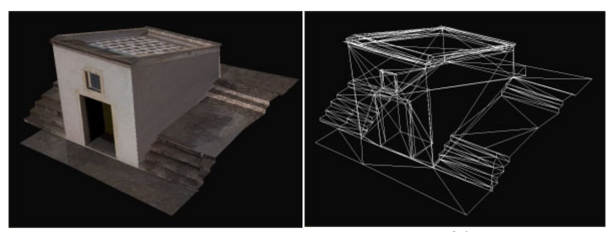
3. TECHNOLOGY USED FOR THE BYZANTINE CRYPT

To model the Byzantine Crypt (Fig.1), we chose a photogrammetric technique for the outside (i.e. main and secondary entrances located above the Crypt) and a laser range scanner that provided plain clouds of 3D points for the Crypt (located underground). Texture information was not available from the scanner and therefore, it was acquired separately with a high-resolution digital camera. Two-dimensional imaging is not only used to record appearance but also to perform geometric measurements and to produce 3D textured models. Proper camera calibration and bundle adjustment algorithms combine in digital photogrammetry to give accurate feature coordinates and reliable pose estimations (Triggs et al., 2000). Many commercial packages perform this task quite nicely (Debevec et al., 1996; El-Hakim, 2001). The model of the main entrance is shown on Figure 2. The model for the secondary entrance was built in a similar way (Figure 3).



a) b)

Figure 2. Crypt main entrance build with photogrammetry, a) textured model, b) wire frame model.

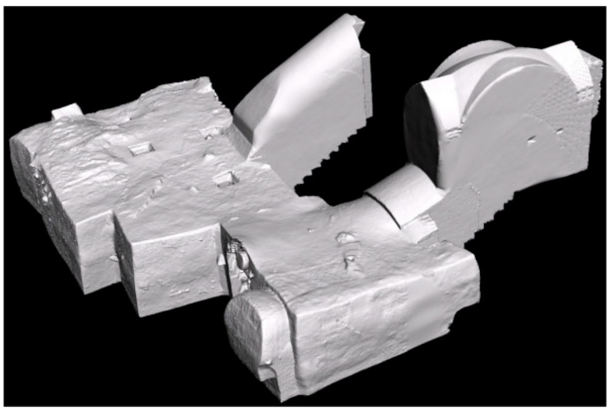


a) b)

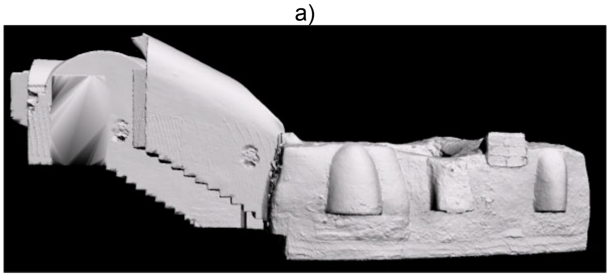
Figure 3. Crypt second entrance build with photogrammetry, a) textured model, b) wire frame of model.

We selected a SLR-type digital camera, the Nikon D1x for the texture acquisition. The CCD sensor has an area of 23.7 x 15.6 mm and an effective pixel count of 4028 x 1324. The output image is re-interpolated to a resolution of 3008 x 1960 pixels (imager ratio 3:2). Both native (NEF) and TIFF formats are available. Proper texturing of the 3D model requires special lighting fixtures in order to control illumination. Good uniformity of the illumination is important in order to ease the processing tasks. In an environment with frescoes, the main problem with lighting is the amount of heat generated by high power lamps. In this case, the amount of heat must be kept to a minimum to avoid damage. Xe flashtubes with a colour temperature of about 5600 K were used. The tubes are UV coated and the stored energy is about 500Ws with duration of 1/700 sec. The manufacturer rates the stability at ±1%. All of the images were acquired with a fixed focal length to ease calibration of intrinsic parameters and an f/22 aperture to produce a large depth of field.

In order to create a dense 3D model of the Byzantine Crypt, a MENSI SOISIC-2000 scanner was used. Table 1 summarizes the specifications of this laser range scanner. This laser scanner can acquire 3D images at a minimal distance of 0.8 m and at up to 10 m with a measurement uncertainty varying between 0.4 mm and 2 mm (distance-dependent). Though the Byzantine Crypt is relatively large (16.5 m by 10 m by 2.5 m), we still wanted to model it with a fairly high spatial resolution. For this size environment, there aren't a lot of range cameras on the market that could provide us with the desired level of spatial resolution and measurement uncertainty. In fact, these distances represent the transition between optical triangulation and time of flight technologies.



a)



b)

Figure 4. Complete 3D model of the Byzantine Crypt shown with synthetic shading, a) view from the outside showing the two entrances of the complete 3D model of the Byzantine Crypt, b) a particular view of the stairs leading to the Crypt. Size 16.5 m × 10 m × 2.5 m, spatial resolution of 5 mm, range uncertainty of 1 mm and accuracy of 15 mm.

In order to keep a quasi-constant spatial sampling on the surface of the walls, 3D vertical scans acquired at 2.5 m were used to build the 3D model. A sampling step of 5 mm was agreed upon in cooperation with an art historian. This gave an average scan time per 3D image of about 80 min. And for that standoff distance, the depth uncertainty was estimated at about 0.8 mm (1 sigma). Figure 4 presents the complete 3D model that would appear if one could see through the ground. From this model, a floor plan was created and is shown on Figure 5. We tested two techniques to align the 3D images, the first based upon spheres positioned strategically in the scene and the second based on data driven alignment (based on ICP) followed by a global alignment. Results are not reported in this paper but demonstrated the advantages of the latter method.

SPECIFICATION	VALUE
Field of View	$46^{o} \times 320^{o}$
Standoff (mm)	800
Maximum range (mm)	10 000
Resolution (X) minimum mesh size	0.1 mm per meter of range
Z measurement Uncertainty-1 σ (mm) Cooperative surface	0.3 @ 800 mm 0.4 @ 2500 mm 0.6 @ 4000 mm
Data Rate (Hz)	100
Scanner size (cm^3)	$73 \times 21 \times 28$
Scanner weight (Kg)	16.3
Output data type	Cloud of points without intensity information

Table 1 SOISIC™ 2000 laser range scanner specifications

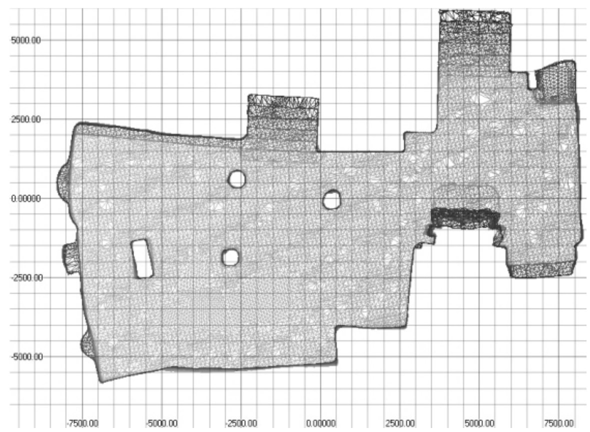


Figure 5. Floor plan generated from an orthographic view of the 3D model of the Crypt showing its dimensions.

4. PUTTING IT ALL TOGETHER: IPT MAPPING ONTO 3D

Projects aimed at the construction of dense 3D and appearance models have become too numerous to be listed here. Each project tries to optimize some part or all of the modeling phases. What we had to deal with was a 3D model that did not have intensity (also know as reflectance channel) data attached to it. This model was created after the merging process (removal of redundancy in overlapped regions) and different resolution models were also created after compressing the polygons to appropriate spatial resolutions. The technique implemented in the commercially available software Polyworks™ is explained in (Soucy and Laurendeau, 1995).

The methodology proposed is very flexible and within reach to non-experts. It uses commercially available software and a small program that combines the re-projection of the 3D points found in the un-textured model file (e.g. VRML format) onto the texture images to give the complete realistic-looking and geometrically correct 3D textured model. This last module will become part of a commercial package. The 2D camera does not have to be rigidly mounted on the 3D camera and therefore 2D images created from digital cameras can be mapped onto the 3D model. These 2D images can be taken specifically for texturing purposes or obtained by other means, e.g., tourist photos, postcard, infrared or ultraviolet images, or even historical photos.

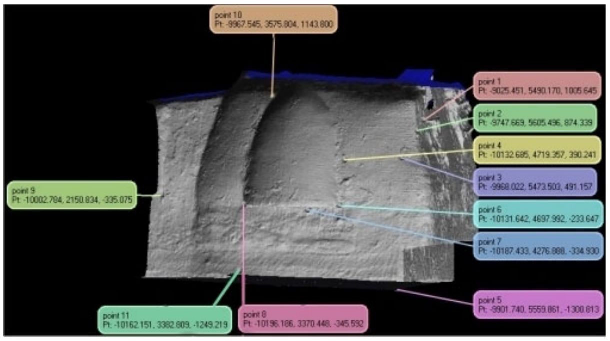


Figure 6. 3D points selection in Polyworks™ using the shaded images

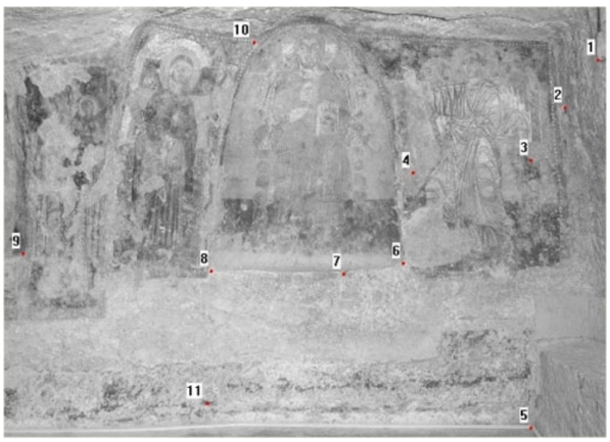
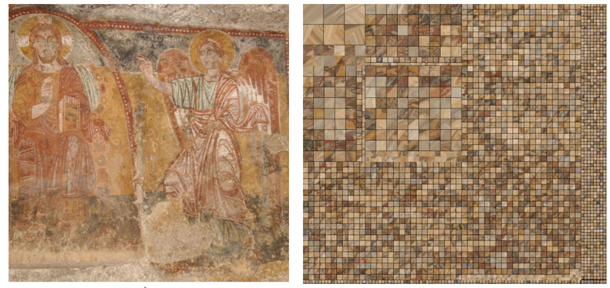


Figure 7. Selection of homologous points in 2D image with ShapeCapture™.

The model is first segmented manually (in the 3D modeling software) into mutually exclusive regions. Each region is mapped onto a region (entirely comprised) that is a subset of one of the 2D images. Then, features are located on a shaded version (see Figure 6) of the 3D image using Polyworks IMinspect™ and an ASCII file is created that contains those 3D points. The same features are located in the 2D image, and the relative position between 2D and 3D cameras is found using the photogrammetric software, ShapeCapture™ (ShapeQuest Inc.) (see Figure 7). We assumed that the 2D camera has already been calibrated and that the 3D points generated by Polyworks™ were imported as control points in ShapeCapture™. Pose estimation in this last software uses the distortion parameters of the lens computed from the camera calibration. We use a 6-parameter lens distortion model. In house software to map texture was then used. This last item will be available in a future release of ShapeCapture™. The calibration can be performed once, before taking the 2D images or if the camera if no longer available, then, 3D data points found on the 3D model can be used for the lens calibration. Other methods based on constraint equations (e.g. perpendicularity or parallelism amongst features) are available, (El-Hakim, 2001). The 2D images can be taken in an angular span of about ± 30 degrees. Grazing angles should be avoided.



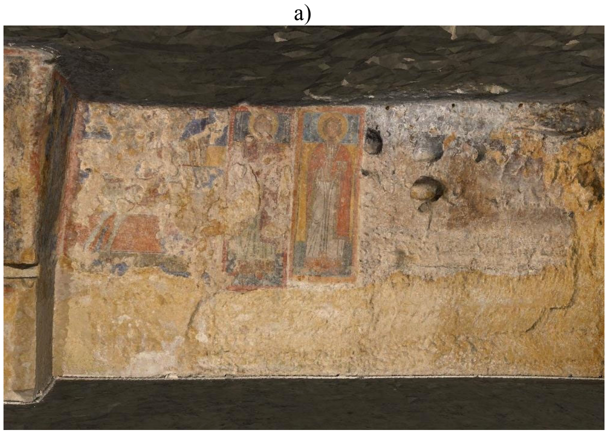
a) b)

Figure 8. Texture mapping methods, a) the preferred way to map texture because it is usable in virtual restoration, b) texture map for the colour per vertex method.

We have experimented with two approaches for the construction of a texture-mapped simplified model, again with the goal of maximizing the use of commercially available software tools.

The <u>first method</u> prepares the data so that it can be entered into the model compression and texture mapping process available in Polyworks™. This technique (Soucy et al., 1996b) requires a triangulated geometric model with a colour value assigned to each vertex. The original high-resolution model is compressed into a simplified model through a vertex removal process. The appearance of the original model is approximated by computing a texture patch for each triangle of the simplified model that approximate the appearance of the area represented by the removed vertices. As part of the geometric compression, the removed vertices are projected onto the larger triangles. When the desired level of compression is reached, a texture image is created. It is a tessellated image where each triangle of the simplified model is mapped onto a colour patch integrating the information from the removed vertices. The method was designed for arbitrary topologies, and possibly incorrect or incomplete models. Thus, it does not attempt to create a piecewise parameterization of adjacent triangles on the model that would be maintained in the texture image. Rather, each triangle is mapped independently (or by adjacent pairs) and after affine transformation of the triangle into an isosceles right angle triangle for efficient packing. Figure 8b illustrates the results. In order to apply this method, the colour information contained in the 2D images must be attached to vertices of the geometric model. Here, the surface sampling of the original 2D images is denser than the geometric sampling. In order to incorporate the colour information in the model, triangles are subdivided to accommodate the new points. The over-sampled model is then fed into the pipeline. One advantage of using this method is that all the texture is embedded in a single, efficiently occupied texture map, and that the algorithm easily allows the generation of maps of different sizes. The major inconvenient of this approach is the requirement to process an excessively large model. But another one is in the usability of the texture map obtained: if there are requirements to modify the images, only global corrections (e.g. contrast, brightness) can be easily applied to the texture. If, in the course of virtual restoration, the original images need to be modified, then the entire compression/texture mapping process must be applied again.

a)



b)

Figure 9. A section of the Crypt, a) synthetic shading replaces one of the colour images, b) the proper colour image mapped with the technique proposed in this paper.

The second method, which we ultimately adopted, simply uses the manually assigned pairings between subsets of 3D triangles and individual 2D images. The only software component required here is a simple mapping program that assigns the texture coordinates to each vertex of the 3D model by solving the collinearity equations used in the photogrammetric module. Because the triangles are relatively small, no perspective correction was applied to the mapped texture triangles. With this method, the original high-resolution 2D images (Figure 8a) are always available for processing. Once modified for a given task, e.g., like virtual restoration, a simple reload of the VRML file in the viewer updates the model. There is no need to re-compute the projection. The realistic looking nature of the model comes from the fact that a calibration of the 2D camera guarantees the geometric quality of the mapping; the mapping uses all the texture data present in the 2D image, without re-mapping like in Figure 8b, and a dense 3D model. Therefore a user can select the level of resolution for an application and then map the texture on a high-resolution 3D model. A section of the Crypt is shown in Figure 9a using synthetic shading after removal of one colour image and in Figure 9b, the corresponding colour image is mapped back onto the surface with the technique proposed above. The matching between 2D and 3D data is performed interactively. We could include in the features detection part of the solution a module that does segmentation and matching between 2D and 3D imagery (Neugebauer and Klein, 1999). We also do not rely on automatic best view computations. If 3D points are occluded then at this point we don't remove the texture mapping around that point. Presently we rely on the ability of the user to take the 2D snapshots and to pick the best point of views when segmenting the 3D model. Another shortcoming is that method does not model reflectance. However some of these features will be added in the near future. Nevertheless, the results are very good especially in a VR room.

Figure 10. View of the texture mapped model of the Crypt from a vantage point that is not physically possible.

a)

b)

Figure 11 Example of a virtual restoration, a) current state of some of the writings, b) enhanced version with some modifications brought to the texture image. Text taken from reference (Fonseca, 1979).

A Byzantine Crypt (rupestrian site - 9th century c.e.), was selected as a way to demonstrate the proposed modeling method. It is characterized by two entrances one leading to the area that served as a cemetery and the other as the church. The church portion is divided into two naves according to a structure that is typical of the period. The Crypt is characterized by a number of frescoes on the walls. One of them, Christ and the Annunciation, is dated at 959 c.e. and is signed by Theophylact. These facts make this Crypt an important heritage site. During the course of history, the floor was lowered in order to make room for a Baroque altar and three pillars replace one that collapsed in the 18^{th} century. Figure 10 presents a view of the texture-mapped model of the Crypt from a vantage point that is not physically possible. In fact, the view is taken outside the Crypt with the texture information map applied on the back of the walls. As an example to illustrate a virtual restoration on textual information, Figure 11a shows the current state of some of the writings on the so-called Theophylact group and Figure 11b presents the enhanced version with some modifications brought to the texture image using the text presented by (Fonseca et al., 1979). These snap shots come from the model.

5. CONCLUSION

Virtual and virtualized environments offer the possibility to expand our abilities in planning, creating and experiencing new surroundings, e.g. virtual tourism on sites that have been closed to the general public or sites not popular as some may want. The resulting visual simulation aims at an exact representation of the real world allowing for photo-realistic rendering, telepresence, and intuitive information queries. People have been trying to improve the sense of realism by using models generated by sensors. If the only goal is the generation of photo-realistic images for visualization, then purely image-based rendering techniques offer a general solution. However, if the goal is to analyze the work, to preserve and share a record of their geometry and appearance, then explicit shape information must be acquired and stored in an adequate representation that also includes some form of appearance modeling.

The potential of modeling as-built reality for heritage applications for such applications as virtual restoration, or as an input to virtualized reality tours was shown with a Byzantine Crypt. A high degree of realism can be attained by those techniques and the context in which the artefacts were discovered or were used can be recreated. Real world acquisition and modeling is now possible. Technological advances are such that difficulties are more of a logistical nature than technological per se. Many techniques exist to digitize small objects with both a high-resolution 3D surface and a view independent surface texture with perfect registration between them. Models of large objects, structures and environments are possible but as demonstrated here require the combination a number of techniques. Many papers in the literature explore both modeling and texture mapping onto dense 3D models but the results are not necessarily accessible to everyone interested in applying this technology. The problem we addressed in this paper is the creation of tools and methods that work with commercial devices and software.

6. ACKNOWLEDGEMENTS

Contributors: E. Bandiera, F. Bergamo, M.C. Catamo, Sac. G. Colavero, D. Lucarella, R. Ingrosso, A. Marra, F. Melcarne, S. Nuccio, P. Pulli, A. Van Den Troost , U. Parrini, M. Biliotti, C. Sempi, V. Blasi, R. De Rinaldis and M. L. Blasi. Co-financing was provided by the *Ministero dell'Istruzione, dell'Università e della Ricerca* and the European Union (F.E.S.R.) for initiative 18 of the *Piano Coordinato delle Università di Catania e Lecce*.

7. REFERENCES

Bandiera E. 1980. *Carpignano Salentino: Centro, Frazione, Casali*, Ed. Lorenzo Capone, Cavallino.

Baribeau, R., Rioux, M., and Godin, G. 1992. Color reflectance modeling using a polychromatic laser range sensor, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), pp. 263-269.

Bernardini, F., Martin, I.M., and Rushmeier H. 2001. High-quality texture reconstruction from multiple scans, *IEEE Transactions on Visualization and Computer Graphics*, 7(4), pp. 318-332.

Debevec, P. E., Taylor, C. J. and Malik, J. 1996. Modeling and rendering architecture from photographs: A hybrid geometry-based and image-based approach. In Proc. of ACM Siggraph 1996, ACM Press/ACM SIGGRAPH, New York. E. Fiume, Ed., Computer Graphics Proc., Annual Conf. Series, ACM.

El-Hakim S.F. 2001. Three-dimensional modeling of complex environments. *SPIE Proceedings vol. 4309, Videometrics and Opt. Methods for 3D Shape Measurement*, San Jose, Jan 20-26.

El-Hakim S.F., Brenner C., and Roth G. 1998. A multi-sensor approach to creating accurate virtual environments, *ISPRS Journal of Photogrammetry & Remote Sensing*, 53, pp. 379–391.

Fonseca, C.D., Bruno, A.R., Ingrosso, V., and Marotta, A. 1979. *Gli insediamenti rupestri medioevali nel Basso Salento*, Ed. Congedo, Galatina, p. 63.

Neugebauer P.J. and Klein K. 1999. Texturing 3D Models of Real World Objects from Multiple Unregistered Photographic Views, In *Eurographics '99 Conference proceedings*, pp. 245-256.

Sequiera, V., Ng, K., Wolfart, E., Goncalves, J. and Hogg, D. 1999. Automated reconstruction of 3D models from real environments. *ISPRS Journal of Photogrammetry & Remote Sensing*, 54, pp. 1-22.

Soucy, M. and Laurendeau, D. 1995. A General Surface Approach to the Integration of a Set of Range Views, In *IEEE PAMI*, 17(4), pp. 344-358.

Soucy, M. Godin, G. Baribeau, R. Blais, F. and Rioux, M. 1996a. Sensors and algorithms for the construction of digital 3-D colour models of real objects, *Proc. of the IEEE ICIP*, Lausanne, Suisse. September 16-19, Vol. II. pp. 409-412.

Soucy, M., Godin, G., and Rioux, M. 1996b. A texture-mapping approach for the compression of colored 3D triangulations, *The Visual Computer*, 12, pp. 503-514.

Stamos I. and Allen P.K. 2000. 3-D model construction using range and image data. *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition. CVPR 2000*, 1, pp. 531-536.

Triggs, W., McLauchlan, P., Hartley, R., and Fitzgibbon, A. 2000. *Bundle Adjustment for Structure from Motion*, Vision Algorithms: Theory and Practice, Springer-Verlag.

SIBA Web Site: http://siba2.unile.it, http://siba3.unile.it/ctle/

Weinhaus M. and Devarjan V. 1997. Texture mapping 3D models of real-world scenes, *ACM Computing Survey*, 29(4), pp. 325-365.