REALITY-BASED 3D DOCUMENTATION OF WORLD HERITAGE SITES:
METHODOLOGIES, PROBLEMS AND EXAMPLES

Fabio Remondino, Alessandro Rizzi
3D Optical Metrology Group, Bruno Kessler Foundation (FBK), Trento, Italy - (remondino, rizziale)@fbk.eu
Commission V, WG V/4

KEY WORDS: Photogrammetry, Laser scanning, Multi-resolution, Multi-sensor, World Heritage Sites

ABSTRACT:
The importance of Cultural Heritage documentation is well recognized and there is an increasing pressure to document and preserve them also digitally. The continuous development of new sensors, data capture methodologies and multi-resolution 3D representations and the improvement of existing ones can contribute significantly to the documentation, conservation and presentation of Cultural Heritage sites and to the growth of research in this field. The article reviews some important documentation requirements and specifications, the actual surveying and modeling methodologies with their limitations and potentialities as well the visualization issues involved in the heritage field. Some examples of UNESCO world heritage sites 3D documentation are reported and discussed.

1. INTRODUCTION

The Heritage sites in the world (natural, cultural or mixed) suffer from wars, natural disasters, weather changes and human negligence. According to UNESCO, a heritage can be seen as an arch between what we inherit and what we leave behind. In the last years, great efforts focused on what we inherit as Cultural Heritage and on their documentation, in particular for visual man-made or natural heritages, which received a lot of attention and benefits from sensor and imaging advances. The importance of Cultural Heritage documentation is well recognized and there is an increasing pressure to document and preserve them also digitally. Therefore 3D data are nowadays a critical component to permanently record the shapes of important objects so that they might be passed down to future generations. This has produced firstly a large number of projects, mainly leaded by research groups, which have realized very good quality and complete digital models (Levoy et al., 2000; Beraldin et al., 2002; Stumpflel et al., 2003; Guidi et al., 2004; Gruen et al., 2005; Guidi et al., 2006; Ikeuchi et al., 2007; El-Hakim et al., 2008; Remondino et al. 2009a) and secondly has alerted the creation of guidelines describing standards for correct and complete documentations.

The actual technologies and methodologies for Cultural Heritage documentation (Ikeuchi and Miyazaki, 2008) allow the generation of very realistic 3D results (in terms of geometry and texture) used for many scopes like archaeological documentation, digital conservation, restoration purposes, VR/CG applications, 3D repositories and catalogs, web geographic systems, visualization purposes, etc. But despite all the possible applications and the constant pressure of international organizations, a systematic and well-judged use of 3D models in the Cultural Heritage field is still not yet employed as a default approach for different reasons: (i) the “high cost” of 3D; (ii) the difficulties in achieving good 3D models by everyone; (iii) the consideration that it is an optional process of interpretation (an additional “aesthetic” factor) and documentation (2D is enough); (iv) the difficulty to integrate 3D worlds with other more standard 2D material. But the availability and use of 3D computer models of heritages opens a wide spectrum of further applications and permits new analyses, studies, interpretations, conservation policies as well as digital preservation and restoration. Thus virtual heritages (Figure 1) should be more and more frequently used due to the great advantages that the digital technologies are giving to the heritage world and to recognize the documentation needs stated in the numerous charters and resolutions.

This contribution reviews some important documentation requirements and specifications, the actual surveying and modeling methodologies with their limitations and potentialities as well the visualization and preservation issues involved in the heritage field. Some examples related to the 3D documentation and modeling of UNESCO World Heritage Sites are also presented and discussed.
2. REALITY-BASED 3D MODELING

“It is essential that the principles guiding the preservation and restoration of ancient buildings should be agreed and be laid down on an international basis, with each country being responsible for applying the plan within the framework of its own culture and traditions” (The Venice Charter, i.e. The International Charter for the Conservation and Restoration of Monuments and Sites, 1964). Even if this was stated more than 40 years ago, the need for a clear, rationale, standardized terminology and methodology, as well as an accepted professional principles and technique for interpretation, presentation, digital documentation and presentation is still evident. Furthermore “(…) Preservation of the digital heritage requires sustained efforts on the part of governments, creators, publishers, relevant industries and heritage institutions. In the face of the current digital divide, it is necessary to reinforce international cooperation and solidarity to enable all countries to ensure creation, dissemination, preservation and continued accessibility of their digital heritage” (UNESCO Charter on the Preservation of the Digital Heritage, 2003). Therefore, although digitally recorded and modeled, our heritage requires more international collaborations and information sharing to digitally preserve them and make them accessible in all the possible forms and to all the possible users and clients.

Nowadays the digital documentation and 3D modeling of Cultural Heritage should always consist of:

• recording and processing of a large amount of three (possibly four) dimensional multi-source, multi-resolution and multi-content information;
• management and conservation of the achieved 3D (4D) models for further applications;
• visualization and presentation of the results to distribute the information to other users allowing data retrieval through the Internet or advanced online databases;
• digital inventories and sharing for education, research, conservation, entertainment, walkthrough or tourism purposes.

The generation of digital 3D models of archaeological objects or sites for documentation and conservation purposes requires a technique with the following proprieties:

• accuracy: precision and reliability are two important factors of the surveying work, unless the work is done for simple and quick visualization;
• portability: the technique should be portable due to accessibility problem of many sites, absence of electricity, location constraints, etc;
• low cost: most archaeological missions have limited budgets and they cannot effort expensive documentation instruments;
• fast acquisition: most sites or excavation areas have limited time of documentation not to disturb works or visitors;
• flexibility: due to the great variety and dimensions of sites and objects, the technique should allow different scales and it should be applicable in any possible condition.

All these properties are often not eligible for a unique technique, therefore most of the surveying projects related to large sites integrate and combines multiple sensors and techniques.

2.1 Techniques

The continuous evolvement of new sensors, data capture methodologies and multi-resolution 3D representations and the improvement of existing ones are contributing significantly to the documentation, conservation and presentation of heritage information and to the growth of research in the Cultural Heritage field. This is also driven by the increasing requests and needs for digital documentation of archaeological sites at different scales and resolutions.

The generation of reality-based 3D models of large and complex sites is nowadays performed using methodologies based on image data (Remondino and El-Hakim, 2006), range data (Blais, 2004; Cignoni and Scopigno, 2008), classical surveying (e.g. total stations or GPS) or graphical and procedural modeling (Mueller et al., 2006). The choice or integration depends on required accuracy, object dimensions, location constraints, system’s portability and usability, surface characteristics, working team experience, project’s budget, final goal, etc.

Although aware of the potentials of the image-based approach and its recent developments in automated and dense image matching (Brown et al., 2003; Goesele et al., 2007; Remondino et al., 2008a), the usability by non-experts and the reliability of optical active sensors (with related range-based modeling software) in certain projects is still much higher, although time consuming and expensive. Active optical sensors (Blais, 2004) like pulsed, phase-shift, triangulation-based laser scanners or stripe projection systems have received in the last years a great attention, also from non-experts, for 3D documentation and modeling applications. Range sensors are easy to use and provide quickly and directly the required 3D data despite their high costs and weights and the usual lack of texture.

On the other hand images require a mathematical formulation to transform the two-dimensional image measurements into three-dimensional coordinates. Image-based modeling techniques (mainly photogrammetry and computer vision) are generally preferred in cases of lost objects, monuments or architectures with regular geometric shapes, low budgets, good experience of the working team, time or location constraints for the data acquisition and processing (Remondino and El-Hakim, 2006). Between the available platforms (space-, airborne and terrestrial), of particular interest are the UAVs (Unmanned Aerial Vehicle), i.e. model helicopters which fly in an autonomous mode, using integrated GPS/INS, stabilizer platform and digital cameras, and which can be used to get images from otherwise hardly accessible areas. Image-based 3D modeling generally requires some user’s interaction in the different steps of the modeling pipeline, reducing its use only to experts. Therefore fully automated methods based on a ‘structure from motion’ approaches (Vergauwen and Van Gool, 2006; Goesele et al., 2007) are getting quite common in the 3D heritage community mainly for visualization and VR applications as not yet fully reliable and precise. Fully automated and accurate approaches were instead presented in (Barazzetti et al., 2009) but the complete automation in image-based modeling is still an open research’s topic.

Many authors (Pomaska, 2001; D’Ayalza and Smars, 2003; English Heritage, 2005) report how the photogrammetric approach allows surveys at different levels and in all possible combinations of object complexities, with high quality requirements, easy usage and manipulation of the final products, few time restrictions, good flexibility and low costs. Different comparisons between photogrammetry and range sensors were also presented in the literature (Böhler, 2005; Remondino et al., 2005; Grussenmeyer et al., 2008). But although many discussions are still opened on which approach and technique is better in which situation, the best actual answer is the combination and integration of the different sensors and techniques.
2.2 Multi-sensor and multi-source data integration

For the 3D documentation of large and complex sites, the state-of-the-art approach (Figure 2) uses and integrates multiple sensors and technologies (photogrammetry, active sensors, topographic surveying, etc.) for the derivation of different geometric levels of detail (LOD) of the scene under investigation. 3D modeling based on multi-scale data and multi-sensors integration is indeed providing the best 3D results in terms of appearance and geometric detail. Each LOD is showing only the necessary information while each technique is used where best suited to exploit its intrinsic modeling advantages.

Since the nineties multiple data sources were integrated for industrial, military and mobile mapping applications. Sensor and data fusion were then applied also in the Cultural Heritage domain, mainly at terrestrial level but in some cases also with satellite, aerial and ground information for a more complete multi-resolution surveys (El-Hakim et al., 2004; Gruen et al., 2005; Rönnholm et al., 2007; Guidi et al., 2009a).

The integration of sensors and data tries to (i) exploit the intrinsic potentials and advantages of each technique, (ii) compensate for the individual weaknesses of each method alone and (iii) achieve more accurate and complete surveying, modeling, interpretation and digital conservation results.

![Figure 2: The 3D multi-sensor and multi-resolution modeling pipeline based on optical sensors and data for the generation of point clouds and textured 3D models.](image)

The multi-resolution concept (Figure 2) should be distinguished between (i) geometric modeling (3D shape acquisition, registration and further processing) where multiple resolutions and sensors are seamlessly combined to model features with the most adequate sampling step and derive different geometric levels of detail (LOD) of the scene under investigation and (ii) appearance modeling (texturing, blending, simplification and rendering) where photo-realistic representations are sought taking into consideration variations in lighting, surface reflectivity, seamless blending of the textures, user’s viewpoint, taking into consideration variations in lighting, surface pentimenti and preparatory drawings. On the left side of the visible spectrum, at shorter wavelengths, there are the UV (Ultraviolet) radiations, still very useful in heritage studies to identify different varnishes and over-paintings, in particular with induced visible fluorescence imaging systems (Pelagotti et al., 2006). All those multi-modal information need to be aligned and often overlapped to the geometric data for information fusion, multispectral analysis or other diagnostic applications (Remondino et al., 2009b).

2.3 Standards in digital 3D documentation

In the last decades many image-based modeling packages and range-based systems came out on the market to allow the digital documentation and 3D modeling of objects or scenes. Many new users are approaching these methodologies and those who are not really familiar with them need clear statements and information to know if a package or system satisfies certain requirements before investing. Therefore technical standards for the 3D imaging field must be created, like those available for the traditional surveying or CMM. A part from standards, comparative data and best practices are also needed, to show not only advantages but also limitations of systems and software. In these respects, the German VDI/VDE 2634 contains acceptance testing and monitoring procedures for evaluating the accuracy of close-range optical 3D vision systems (particularly for full-frame range cameras and single scan). The American Society for Testing and Materials (ASTM) with its E57 standards committee is trying to develop standards for 3D imaging systems for applications like surveying, preservation, construction, etc. The International Association for Pattern Recognition (IAPR) created the Technical Committee 19 - Computer Vision for Cultural Heritage Applications - with the goal of promoting Computer Vision Applications in Cultural Heritage and their integration in all aspects of IAPR activities. TC19 aims at stimulating the development of components (both hardware and software) that can be used by researchers in Cultural Heritage like archaeologists, art historians, curators and institutions like universities, museums and research organizations. As far as the presentation and visualization of the achieved 3D models concerns, the London Charter ([http://www.londoncharter.org/](http://www.londoncharter.org/)) is seeking to define the basic objectives and principles for the use of 3D visualisation methods in relation to intellectual integrity, reliability, transparency, documentation, standards, sustainability and access of Cultural Heritage.

3. PROBLEMS AND BOTTLENECKS

The actual problems and main challenges in the 3D surveying of large and complex sites or objects arise in every phase, from the data acquisition to the visualisation of the achieved 3D results. Selecting the appropriate methodology (sensor, hardware, software), the appropriate modeling procedures, designing the production workflow, assuring that the final result is in accordance with all the given technical specifications and being able to fluently display and interact with the achieved 3D model are the actual great challenges.

3.1 Data acquisition

In case of satellite and aerial images, the availability of the data could be a problem due to weather conditions or impossibility to flight due to restrictions. For terrestrial acquisition, size, location and surface (geometry and material) of the object or site can create several problems. The dimensions and
accessibility problems (due to location, obstructions, rough or sloped terrain with stones, rocks and holes, unfavourable weather conditions, etc.) can cause delays, occlusions and can result in missing sections or enforce wide-baseline images and poor geometric configurations. The complexity of some parts can create self-occlusions or holes in the coverage, in addition to the occlusions from plants, trees, restoration scaffolds or tourists. The absence of high platforms for a higher location of the data acquisition might cause missing parts, e.g. for the roofs. For laser scanning, the object material (e.g. marble) has also an important influence on the acquired data as can cause penetration effects (Godin et al., 2001; Lichti and Harvey, 2002; Guidi et al., 2009b). In those projects dealing with multi-sensor integration, problems arise from the seamless and correct geometric integration of the acquired data.

3.2 Data processing and point cloud generation

For image-based approaches, terrestrial digital cameras must be accurately calibrated, preferably in a controlled lab environment, with a 3D testfield and a bundle adjustment solution with additional parameters to fully compensate for systematic errors (Remondino and Fraser, 2006). As no commercial procedure is already available for automated markerless tie point extraction from terrestrial convergent images, the camera orientation phase is still highly interactive, although some recent work seems to be promising (Barazzetti et al., 2009). For the surface measurement, manual and semi-automated measurements are still much more reliable in case of complex architectural objects. For small areas or ornaments rich of details, dense matching or shape-from-X techniques can be instead selected to derive dense 3D point clouds (Remondino et al., 2008a).

As far as range-based approaches concerns, the first operations performed on the acquired data are possible errors and outliers removal, noise reduction and holes filling (Weyrich et al. 2004), followed by the alignment (or registration) of the multiple scans. The registration phase is quite straightforward although the identification of homologues points between the overlapping point clouds is still fully interactive.

In the data acquisition and processing of large sites, in particular for those digitized with active optical sensors, we should consider that:
- the huge amount of (range) data makes very time consuming and difficult their processing at high resolution, yet processing at low resolution creates accuracy problems;
- combining data acquired with different sensors, at different resolution and viewpoints can affect the overall accuracy of the entire 3D model if not properly considered;
- despite combining several sensors, some gaps and holes can still be present in the model, requiring filled and interpolated surface patches not to leave them visible and unpleasant;
- the used sampled distance in scanning is rarely optimal for the entire site or object, producing under-sampled regions where edges and high curvature surfaces are present and over-sampled regions where flat areas are.

3.3 3D modeling

Once a point cloud (i.e. unstructured data) is available, a polygonal model (i.e. structured data) is usually generated to produce the best digital representation of the surveyed object or scene. Some repairing to close holes, fix incorrect faces or non-manifold parts are always necessary and time consuming. Those errors are visually unpleasant, might cause lighting blemishes due to the incorrect normals and the computer model will also be unsuitable for reverse engineering or physical replicas. Moreover over-sampled areas should be simplified while under-sampled regions should be subdivided.

Finally photo-realism, defined as having no difference between a view rendered from the model and a photograph taken from the same viewpoint, is generally required and obtained with the texture mapping phase. Generally problems rise firstly from the time-consuming image-to-geometry registration and then come because of variations in lighting, surface specularity and camera settings. Often the images are exposed with the illumination at imaging time but it may need to be replaced by illumination consistent with the rendering point of view and the reflectance properties (BRDF) of the object (Lensch et al., 2003). High dynamic range (HDR) images might also be acquired to recover all scene details (Reinhard et al., 2005) while color discontinuities and aliasing effects must be removed (Debevec et al., 2004; Umeda et al., 2005; Callieri et al., 2008).

3.4 Realistic visualisation of the 3D results

The ability to easily interact with a huge 3D model is a continuing and increasing problem. Indeed model sizes (both in geometry and texture) are increasing at faster rate than computer hardware advances and this limits the possibilities of interactive and real-time visualization of the 3D results. Due to the generally large amount of data and its complexity, the rendering of large 3D models is done with a multi-resolution approach displaying the large meshes with different levels of detail and simplification approaches (Luebke et al., 2002; Cignoni et al., 2005; Dietrich et al., 2007).

4. EXAMPLES

4.1 The Etruscan Necropolis of Tarquinia (Italy)

Together with Cerveteri, these are the two large Etruscan cemeteries with different types of burial practices from the 9th to the 1st century BC, bearing witness to the achievements of Etruscan culture which over nine centuries developed the earliest urban civilization in the northern Mediterranean. Some of the tombs are monumental, cut in rock and topped by impressive tumuli (burial mounds). Many feature carvings on their walls, others have wall paintings of outstanding quality. The necropolis near Cerveteri, known as Banditaccia, contains thousands of tombs organized in a city-like plan, with streets, small squares and neighbourhoods. The site contains very different types of tombs: trenches cut in rock; tumuli; and some, also carved in rock, in the shape of huts or houses with a wealth of structural details. These provide the only surviving evidence of Etruscan residential architecture. The necropolis of Tarquinia, also known as Monterozzi, contains 6,000 graves cut in the rock. It is famous for its 200 painted tombs, the earliest of which date from the 7th century BC (http://whc.unesco.org).

A multi-resolution and multi-modal 3D model of one important grave in Tarquinia (“Caccia e Pesca”, composed of 2 rooms of spanning approximately 5 x 5 x 2 m each one) was realized as a pilot project for the “Soprintendenza per i Beni Archeologici dell’ Etruria Meridionale” (SBAEM) and in collaboration with the company Ari-Test (http://www.art-test.com) which was responsible for the multispectral data acquisition.

A surveying strategy similar to Figure 2 was employed, choosing the most appropriate method in relationship to the research goals and object’s scale. The acquired data comprise:
- geometric data: a TOF scanner surveying acquired a massive amount of range data for the exterior (10 stations @ 1 cm geometric resolution, ca 2 Mil. points) and the underground interior (13 stations @ 4 mm sampling step, ca 14 Mil. points). After the geometric alignment and data reduction a complete mesh was produced for further rendering and interactive visualization purposes.

- appearance data, constituted of:
  (i) visible images for texturing purposes: for the photo-realistic rendering of the final 3D model, ca 160 HDR textures were acquired with a 13.5 Mpixel Kodak DCS camera pre-calibrated in the lab at a focal length setting of 50 mm. A constant illumination in the underground rooms was achieved using cold neon lights (to avoid heating effects on the frescos) and a spot-meter.
  (ii) multispectral images for diagnostics studies: on some selected areas, visible reflectance, IR reflectography and UV induced fluorescence images were acquired using interferential filters in front of a calibrated cooled CCD camera. Those images were afterwards calibrated, registered, processed and overlapped onto the 3D geometry to perform quantitative analysis and differentiate pigments, being them present, hidden or disappeared to the naked eye. Indeed all the materials having the same color in a certain light have different chemical compositions and reflectance spectra. The integration between the different data sources was essential to overcome some limits of each method and to have as result the complete geometric and appearance information about materials and techniques used to build the heritage.

In Figure 3 some results of the geometric and appearance modeling are reported with also examples of the multispectral data. The final photo-realistic 3D model with its multispectral layers is now a fundamental basis for conservation and restoration policies to help the local superintendence in preserving the vanishing of the frescoes and correctly restoring the damaged areas.

4.2 Pompeii and its Roman Forum

When Vesuvius erupted on 24 August AD 79, it engulfed the two flourishing Roman towns of Pompei and Herculaneum, as well as the many wealthy villas in the area. These have been progressively excavated and made accessible to the public since the mid-18th century. The vast expanse of the commercial town of Pompei contrasts with the smaller but better-preserved remains of the holiday resort of Herculaneum, while the superb
wall paintings of the Villa Oplontis at Torre Annunziata give a vivid impression of the opulent lifestyle enjoyed by the wealthier citizens of the Early Roman Empire (http://whc.unesco.org).

The large and complex Roman Forum in Pompeii (approximately 150 x 80 m large with more than 300 scattered archaeological finds on the ground) was digitally reconstructed (Figure 4) integrating aerial images, terrestrial laser scanning and close-range images (Guidi et al., 2009a). The geometric resolution of the 3D data spans from some cm (from the aerial views) down to few mm (3D models of relieves derived with terrestrial photogrammetry), with an intermediate LOD given by the TOF range data.

Figure 4: The entire roman forum in Pompeii (ca 180 x 50 m with more than 300 scattered finds on the ground) digitally reconstructed integrating photogrammetry and TOF scanning.

Some objects of particular archaeological interest were geometrically modeled in high-resolution (Figure 5) by means of advanced image matching (Remondino et al., 2008a).

Figure 5: A detailed ornament automatically reconstructed using the CLORAMA image matching software (4DiXplorer AG).

The entire 3D model, since it is geo-referenced, can be easily linked to existing archaeological databases (Manferdini et al., 2008, Figure 6). The database - 3D model relationship can be implemented in two ways: (i) from the geometrical model to the connected data, for explaining historical and conservation details of a specific artifact in the Forum and (ii) from a specific document or philological detail to the corresponding location in the 3D space.

4.3 Angkor Wat (Cambodia) and its temples and statues

Angkor is one of the most important archaeological sites in South-East Asia. Stretching over some 400 km², including forested area, Angkor Archaeological Park contains the magnificent remains of the different capitals of the Khmer Empire, from the 9th to the 15th century. They include the famous Temple of Angkor Wat and, at Angkor Thom, the Bayon Temple with its countless sculptural decorations. UNESCO has set up a wide-ranging programme to safeguard this symbolic site and its surroundings (http://whc.unesco.org). Some 3D modeling results of the archaeological temples and Buddha figures are reported in (Gruen et al., 2001; Sonnemann et al., 2006; Ikeuchi et al., 2007) and Figure 7.

Figure 6: Example of 3D modeling and semantic classification of an archaeological find in Pompeii for the connection between the reconstructed 3D shape and the information contained in a database.

Figure 7: Angkor Wat and Phnom Bakheng temples (respectively) reconstructed from aerial images using stereo-measurements.

5. CONCLUSIONS AND OUTLOOK

The article reviewed the actual surveying and 3D modeling methodologies with their limitations and potentialities as well the visualization issues involved in the heritage field. Despite the fact that the 3D documentation is not yet the state-of-the-art in the heritage field, the reported examples show the potentialities of the modern technologies to digitally document and preserve our heritages as well as share and manage them. Further examples of UNESCO World Heritage Sites digitally reconstructed and modeled are Mount Everest (Gruen and Murai, 2002), the Bamiyan valley, Afghanistan (Gruen et al., 2004, 2005; Gruen and Hanusch, 2008), the Geoglyphs of Nasca, Peru (Lambers and Gruen, 2003, Lambers et al., 2004), the Acropolis of Athens, Greece (Moullou et al., 2008; El-Hakim et al., 2008; Remondino et al., 2008b; Tsingas et al., 2008) just to mention some works.

References from Journals:


References from Books:


References from Other Literature:


Acknowledgements

Acknowledgements are due in particular to Prof. A. Gruen and Dr. M. Sauerbier (ETH Zurich, Switzerland), Prof. G. Guidi and Dr. M. Russo (Politecnico of Milan, Italy), Dr. S. El-Hakim and Dr. A.J. Beraldin (NRC Canada), S. Girardi (FBK Trento, Italy), A. Pelagotti (INOA-CNIR, Italy), Dr. A.M. Manferdini and Prof. M. Gaiani (Univ. of Bologna, Italy) co-authors in some publications, researches and field works for the reported projects and examples.