

3D DIGITIZATION OF AN HERITAGE MASTERPIECE - A CRITICAL ANALYSIS ON QUALITY ASSESSMENT

F. Menna ^{a,*}, E. Nocerino ^a, F. Remondino ^a, M. Dellepiane ^b, M. Callieri ^b, R. Scopigno ^b

^a 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Email: (fmenna, nocerino, remondino)@fbk.eu, Web: <http://3dom.fbk.eu>

^b Visual Computing Laboratory, ISTI-CNR, Pisa, Italy
Email: (dellepiane, callieri, scopigno)@isti.cnr.it, Web:<http://vcg.isti.cnr.it>

Commission V, WG V/4

KEY WORDS: Accuracy, Photogrammetry, Laser scanning, Deformation analysis, Camera Network, ICP

ABSTRACT:

Despite being perceived as interchangeable when properly applied, close-range photogrammetry and range imaging have both their pros and limitations that can be overcome using suitable procedures. Even if the two techniques have been frequently cross-compared, critical analysis discussing all sub-phases of a complex digitization project are quite rare. Comparisons taking into account the digitization of a cultural masterpiece, such as the Etruscan Sarcophagus of the Spouses (Figure 1) discussed in this paper, are even less common. The final 3D model of the Sarcophagus shows impressive spatial and texture resolution, in the order of tenths of millimetre for both digitization techniques, making it a large 3D digital model even though the physical size of the artwork is quite limited. The paper presents the survey of the Sarcophagus, a late 6th century BC Etruscan anthropoid Sarcophagus. Photogrammetry and laser scanning were used for its 3D digitization in two different times only few days apart from each other. The very short time available for the digitization was a crucial constraint for the surveying operations (due to constraints imposed by the museum curators). Despite very high-resolution and detailed 3D models have been produced, a metric comparison between the two models shows intrinsic limitations of each technique that should be overcome through suitable onsite metric verification procedures as well as a proper processing workflow.

1. INTRODUCTION

Several reasons may turn some Cultural Heritage (CH) assets in unique witnesses of past human civilizations, gaining popularity for their beauty and eventually becoming renowned all over the world as masterpieces. Their uniqueness and vulnerability at the same time put them in a priority position in the list of all the heritage assets that still need to be converted in digital form. This requires the designing and implementation of proper procedures for a correct, accurate and reliable 3D digitization of the asset. As most of CH exhibits and finds, most masterpieces are stored or put on display in locations that are far from being ideal for their 3D digitization. Once in a while they are freed from the usual position for restoration purposes or alike, offering an irrefutable opportunity for 3D digitization activities. Unfortunately, restoration facilities are not always the best places for carrying out a 3D survey, offering narrow spaces and uncontrolled light environments, among many other issues.

The combination of all these factors often makes the 3D digitization of masterpieces such as the Sarcophagus of the Spouses (Fig. 1) a demanding task. High flexibility and portability of the techniques as well as highly skilled and professional operators are needed for accomplishing the digitization in a constrained time. Redundancy, real-time pre-processing as well as onsite verifications and checklists are some of the tasks to be accomplished by the digitization team, usually under time pressure.

This paper wants to critically highlight the issues and operative decisions taking place in a 3D digitization project of a valuable masterpiece such as the Sarcophagus of the Spouses. We aim at clearly defining which are the most critical steps, not only the technical ones but also those related to the safety of the asset itself. Indeed, a 3D digitization project should always put safety at first, both for the surveying personnel and of course for the CH asset.



Figure 1. The Sarcophagus of the Spouses in its temporary location in Villa Giulia Museum, Rome.

* Corresponding author.

The Digital Michelangelo project (Levoy et al., 2000) is an early comprehensive example in this sense providing critical insights on how to reduce such risks in 3D scanning projects. The 3D Optical Metrology unit (<http://3dom.fbk.eu>) of FBK Trento and the Visual Computing Lab (<http://vcg.isti.cnr.it>) of CNR-ISTI Pisa were involved in the data acquisition and processing of the Sarcophagus, the former using close-range photogrammetry whereas the latter using a triangulation-based laser scanner. While a generalized comparison is not easy to obtain, the paper presents a critical overview of the two digitization processes, discussing the planning phase, the data acquisition and the subsequent processing steps.

1.1 The artwork and project aims

The Sarcophagus of the Spouses is a late 6th century BC Etruscan anthropoid Sarcophagus. It is made of terracotta and was once brightly painted. It depicts a married couple reclining at a banquet together in the afterlife. It was found during the 19th century excavations in the Cerveteri necropolis, Italy. The Sarcophagus was fragmented in hundred pieces, which required a complex restoration in 1893. The idea of performing a 3D digitization of the Sarcophagus was born in the early 2013 when a unique multimedia exhibition was planned to take place in Italy in two different locations in 2014-2015: Palazzo Pepoli in Bologna and the Etruscan Museum of Villa Giulia in Rome. The theme of the exhibition was "*The Etruscan journey towards the afterlife*", narrated through a selected collection of objects, all of special value, preserved and displayed at the Museum of Villa Giulia and then exhibited in the History Museum of Bologna. Despite being digitization activities mostly committed to the multimedia exhibition project (Russo et al., 2013), archaeologists were also interested in the creation of a faithful digital reproduction of the Sarcophagus. A detailed and precise digital acquisition and 3D modelling of the Sarcophagus was thus decided, aimed at using different technologies also to benchmark the existing 3D recording and modelling techniques (Remondino et al., 2013) and to produce best practices and guidelines.

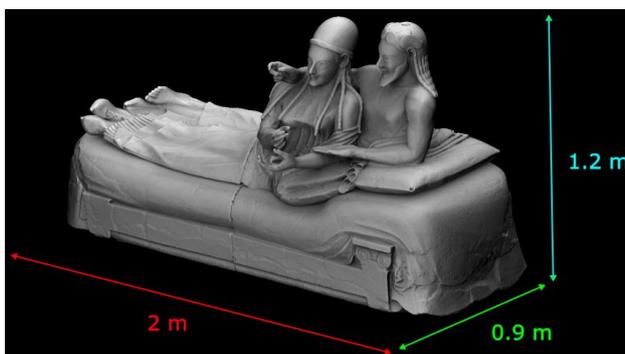


Figure 2. Sarcophagus encumbrance without the supporting legs.

1.2 Deliverables, requirements and constraints for the Sarcophagus project

In agreement with the curator, a 1:1 scale was chosen for the photographic texture while a 0.5 mm spatial resolution for the 3D model was considered sufficient to their aims. Obviously, lighter and simplified 3D models can be easily obtained from the highest resolution ones applying shape-optimized subsampling methods. The Sarcophagus is normally on display at the National Etruscan Museum of Villa Giulia in Rome (Italy) where it is preserved in a glass showcase along the exhibition itinerary. For this project the Sarcophagus was placed on a soft wooden stage covered by a

carpet (Fig. 1). The available free space around the Sarcophagus was not more than 2-2.5 meters, and half a meter of this space was occupied by the stage. Additionally, other neighbouring exhibits in their own showcase were present around the Sarcophagus needing additional care during the survey. The Sarcophagus dimensions without the supporting legs are reported in Figure 2. Only two original legs (the ones on its front side) are preserved and are today attached with magnets to the supporting frame adding approximately 30 cm to its height. The height of the Sarcophagus is approximately 1.4 m with legs thus it forced both surveying teams to use a step ladder and a very tall tripod in order to properly digitize the top of the artefact heads. The time at disposal for the teams was no more than a working day (8 hours) plus the successive morning (5 hours). Two windows covered with white diffusing curtains were next the two ends of the Sarcophagus. Artificial lights from the ceiling and nearby display cases could be switched off when needed.

2. DATA ACQUISITION PLANNING

Acquisition planning is a key step in a 3D reconstruction process both for laser scanning and photogrammetry. Planning a survey must take into account several factors: (i) project requirements (geometry resolution and final texture quality/resolution, purpose of the 3D product, etc.), (ii) object characteristics (dimension, material and surface characteristics, etc.), (iii) environment and time constraints, (iv) budget and (v) characteristics of the selected technical equipment.

For an image viewed at near distance of distinct vision (≈ 250 mm), any subject detail resolved and recorded optically within an image circle of 0.2-0.3 mm diameter may not be perceived (Ray, 2002). Therefore a 1:1 scale photographic texture requires details as fine as 0.2-0.3 mm to be visible on the final texture, a requirement that may arise depth of field (DOF) issues in close-up regime imaging.

To take on the challenging requirements and safety constraints of the project, an accurate analysis of the scenario was undertaken. Meticulous and above all safe positioning of the photographic camera and 3D scanner are needed given the vulnerability of the asset. The first preventive measures considered during the planning were to avoid overhanging objects as much as possible on the Sarcophagus, counterbalancing with weights on the tripods legs and watching and holding them during most delicate operations. A team of two people was judged the minimum for safe management of the 3D digitization in the museum. Working hard on repetitive tasks for hours can be stressful and introduce errors, therefore operators with interchangeable skills were selected.

Although there is a lack of painted texture, the Sarcophagus' surface and material (terracotta) are quite cooperative both for photogrammetry and laser scanning, showing many small features that, if properly lit and imaged, can be used for automatic photogrammetric measurements. However, in some areas, the artefact's surface due to remains of the original painting, restoration and conservation treatments, results semi-glossy under slanted angles, thus demanding a proper lighting set up to avoid specular reflections.

The acquisition planning finally revealed two main *blocks* that require different acquisition strategies: (i) the lower and largest part, i.e. the couch, characterised, excluding the feet, by low-frequency geometrical features; (ii) the two figures which present higher geometrical detail and complexity, with many occluded parts.

2.1 Laser scanning planning

Given the fine geometric details on the artefact and the resolution required, we chose to use a triangulation-based scanning device. Active triangulation scanners generally have multiple lenses sets or camera settings, to cover areas of different sizes (with different resolutions). These area/resolution setups are just a finite number and the selection depends on the characteristics of the object, on the sampling resolution needed, and the practical constraints of the working setup. A smaller area covered in each scan means a higher sampling resolution, but also a more fragmented digitization (i.e. more acquisitions), which requires more time to be sampled and more alignment work. Conversely, a wider area covered in each scan produce a lower sampling resolution, but the surface coverage is quicker, and the alignment possibly faster. Following these considerations and given the short time available to carry out the digitization, the CNR-ISTI team employed a Konica-Minolta Vivid 910, a triangulation-based laser scanner (Fig. 3a and b). Two different lens sets were used: *wide* (covering ca 50x40 cm in each shot) for the less detailed and lower area of the Sarcophagus and *medium* (ca 30x20 cm shot coverage) for the details and human figures. Even though it was not a very “recent” instrument, it was still the most versatile device in its category and it was able to work at the required resolution. During the planning, possible arising problems were also considered:

- *attainment of a rigid alignment*: range scans obtained with a triangulation scanner are generally aligned only using geometric redundancy of overlapping scans. Since the lower part of the Sarcophagus body is pretty featureless, we feared that a good alignment of scans might not be possible in case of small coverage/overlap. On the contrary, if each scan covers a wider area, this would result in a more stable and rigid alignment since the overlapping regions would have been larger. On the other hand, the human figures have plenty of geometrical features, making the alignment easier and certainly more rigid.

- *mixing of various resolutions*: while the aforementioned choice is sound considering time constraints and registration issues, the mixing of two large chunks of data digitized at a different resolution requires some care, as the alignment between the two parts may be tricky and the final merging may still show a different “roughness” of the 3D surface. However, we evaluated that the ratio between the two resolution sampling densities was below 1:2, therefore a still manageable level.

- *possible global deformation* of the final reconstructed model: if the range scans are very small with respect of the total size of the artwork, the alignment error accumulates among the range scans. This could create some degree of deformation of the reconstructed mesh. Ideally, the larger and fewer the scans are, the more rigid is the final alignment. In a normal situation, to prevent warping and deformations, external rigid references should have been placed and used in the alignment phase, e.g.: using external markers, or having a single scan covering the whole object, taken for example with a terrestrial laser scanner (Callieri et al., 2011). In this case, however, time constraints prevented such strategy.

2.2 Photogrammetric planning

From a photogrammetric point of view, image acquisition planning encloses camera network design (Fraser, 1996) as well as rough predictions on photographic settings to use, such as exposure time, sensitivity settings ISO, aperture value and hence DOF. This permits to select the appropriate lighting power, diffuser, stands etc.

A safe distance from the object and the necessary ground sample distance (GSD for the image texture spatial resolution) ruled the

camera network planning. A digital single-lens reflex (DSLR) Nikon D3X 24 megapixel camera featuring a 6 micron pixel size, coupled with a Nikkor 50 mm f/1.8 D lens, was chosen for the acquisition of 14-bit images. Focused at 1 m and using the nominal value of the focal length, the GSD resulted 0.12 mm on the sharp focus plane (1m). The main body of the Sarcophagus, i.e. the couch, was then planned to be acquired with an aerial-like camera network consisting of several strips overlapping some 80% along- and 60% across-track, keeping the camera optical axis orthogonal to a hypothetical average object surface within the field of view. Additionally, rolled and convergent images were planned to strengthen the geometry and help getting more accurate calibration parameters during self-calibration (Nocerino et al., 2014). The approach we generally use is similar to the one shown in Alsadik et al. (2013), where a virtual cage of cameras can be thought around the object and then specific camera positions excluded if not satisfying certain criteria. In general, only images outside the convex hull of the asset and its ground encumbrance were considered.

For highly curled or recessed parts that stretch in a specific direction, such as the garments and braids, images with strips parallel to that direction were planned in order to see the same part inside the curled area in at least two images.

As far as lighting is concerned (Fig. 3d and e), a rigorous approach should take into account several factors such as light positions, their size and orientations for each photograph. To avoid casting shadows visible in the images, the source of light should originate exactly from the centre of perspective of the objective lens, i.e. an impossible condition to be realized in practice. Therefore, lights must be positioned symmetrically with respect to the optical axis of the camera. Under this condition the shadows projected by a lamp or strobe can be filled from the other symmetrical source of light. Diffuse light works similarly, being the source of light virtually everywhere around the object. Furthermore cross polarization of the light was used for the acquisition. Cross polarization is a very diffuse technique in the field of optical microscopes where polarized light is used as a contrast-enhancing technique. On the contrary it is not very popular in photogrammetry and related disciplines as it requires a controlled light environment adding also some extra time for a proper tuning of light and filters for each photographs. As advantage it gives images virtually free from specular reflection components leaving thus only the diffuse one. Colours are more saturated (pure) and lighting more homogeneous all over the photographic subject.

3. ACQUISITIONS

When planning the time needed for data acquisition, temperature acclimatization of the equipment must be considered as it may take some hours for the scale bars, scanners and cameras to reach the ambient temperature.

Calibration of triangulation scanners is often a mandatory procedure to be accomplished before the survey starts. Some scanners just have a self-calibration phase or do not require it at all, due to the mechanical stability of their components.

The need of camera calibration for photogrammetric measurements (Remondino and Fraser, 2006) is today more and more discussed. If a good camera network can be achieved, self-calibration should not differ from calibration on a fixture with targets. Nevertheless it is still broadly accepted as good practice to perform at least one calibration just before and/or after the survey to check the camera calibration parameters have not changed significantly. In case of restricted acquisition time, it is better to have real-time feedback to make sure everything is



Figure 3. Data acquisition: the adopted triangulation-based laser scanner, a Konica-Minolta Vivid 910 (a, b); Gretag Macbeth colour checker (c); photogrammetric survey with artificial light and camera mounted on a tripod (d,e,f).

going smoothly as planned. For laser scanning this may be accomplished by checking and aligning each scan right after they are acquired. With old generation hardware/software, this required one of the surveying personnel to be dedicated to this task, whereas today many scanning software provide real-time feedback on data quality and on-the-fly alignment. From a photogrammetric point of view, a real-time image checking and processing could be achieved working in tethering mode (camera connected to a laptop) or by removing the memory cards or using in-camera Wi-Fi/Bluetooth capabilities or Wi-Fi memory cards. Once the images are on a laptop, an adjustment procedure can be performed on low-resolution images to check the acquired data.

3.1 Laser scanning acquisition

The surface of the artefact was not particularly difficult for the scanner. However, some areas were a bit shinier and some other slightly darker than the rest, requiring a tuning of the scanner parameters. The scanning took 5 hours, producing 66 scans acquired with the wide-area lens and 285 scans acquired with the medium-area lens. The digitization strategy for this kind of artworks is pretty straightforward (following standard, common sense, considerations): the scanner is moved around the artefact, trying to keep a somehow constant distance from the subject (it depends on the device and lenses used, in this case, around 70-

80cm), a straight view-direction with respect to the surface and enough overlap among the range maps. The reason is twofold: (i) a sufficient overlap is necessary in the alignment step, as alignment rely on the geometry of the common area between scans; (ii) enough overlap means that, probably, most of the subject's surface is digitized from more than one direction and this may reduce the sampling noise when merging the data notwithstanding an increase of the overall sampling density. Since obtaining a direct feedback on the degree of overlap from the digitization software is not always possible, the correct amount of overlap must be evaluated by the operator, at the moment of deciding the next shoot. As a rule of thumb, overlap should never be less than 25%, and a good interval is between 30% and 50% (for difficult areas).

When switching from one lens set to the other, great care was posed in order to have a pretty large overlap between the two areas, to ensure a good alignment between the two datasets. Figures 4a and b show the area covered by the wide and medium lenses. A photographic acquisition was also required for texture mapping purposes and it was carried out in a couple of hours of the next morning.

Since the light sources we had could not completely overcome the existing lighting coming from two opposing windows, we decide to exploit it, and just try to make it more uniform.

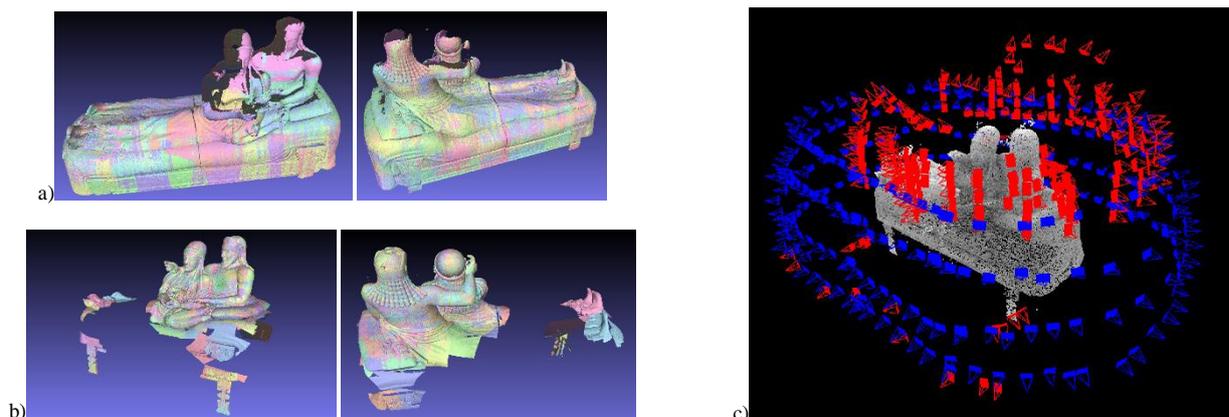


Figure 4. Laser scans acquired with the wide (a) and medium lenses (b). Photogrammetric camera network (c): images acquired for the main block are displayed in blue, images acquired for the most difficult parts (spouses' trunks, hands, feet) are in red.

As the light coming from the windows was already “diffuse” and not “direct” (a direct, strong lighting is not suitable for texture mapping), we used the camera-mounted flash directing it towards the ceiling. This bounced light, added to the ambient lighting, produced a quite smooth illumination. All photos were radiometrically calibrated using a Gretag Macbeth colour checker (Fig. 3c).

3.2 Photogrammetric acquisition

Given the low budget and strict timing at disposal for the project, only a lightweight photographic gear was selected. The whole gear fit in two regular suitcases that two people carried on public transportations. The 3DOM-FBK team was allowed to leave the photogrammetric gear in the Etruscan museum of Villa Giulia the day before the survey, for allowing equipment acclimatization. Before starting the survey some tests were carried out on sample surfaces that were considered more problematic due to their glossy appearance or lack of texture.

Polarized filters were mounted on both lamps and camera. Figure 5 shows the effect of an image acquired with and without cross polarization. The improvement provided to image quality by increasing the contrast and removing specular reflection is evident. Artificial lights in the room were switched off and the influence of ambient light was evaluated shooting with and without the photographic lamps switched on. In our case the ambient light contribution was mostly non influential and when it was the case due to sunlight entering the windows it was shaded with a panel. The exposure was measured on a grey card, manually set to 1 sec at f/16 ISO 400 and kept almost for the whole acquisition. The computed DOF using a circle of confusion of 3 pixels (corresponding to a GSD of about 0.3 mm) was some 22 cm. Throughout the survey the distance of 1m from the object was checked both using a measuring tape and the in camera confirmation focus function in the viewfinder.



Figure 5. A test image of the hands and a scale bar acquired with (left) and without (right) cross polarization.

As mentioned, for the lower part of the artefact, an aerial-like camera network was used. After each strip, one operator set up the tripods and lights at the new height while the other one downloaded the images and ran a photogrammetric adjustment on the collected data. The onsite processing and analyses were performed on a Dell Precision M4600 portable workstation. The overall acquired camera network is shown in Figure 4c, where the images belonging to the main block are coloured in blue whereas the images depicting the details and figures are presented in red.

In order to scale the photogrammetric results, being a small artefact with no possibility to use ground control points (GCP), reference scale bars were adopted. It is generally advisable to use more than a single scale bar or to move it in the measurement volume to have a redundant check. For the Sarcophagus survey, only one lightweight aluminium scale bar, 316.5 mm long, was employed (Figure 5). The size was decided to facilitate its transportation and, above all, to move it in an easier and safer way close to the masterpiece. The scale bar was imaged in different positions all around the Sarcophagus, taking care of acquiring it at least once in each image group. The whole survey took approximately 13 hours in two days (8+5 hours) ending up with about 520 photographs. A camera calibration was carried out at the end of the survey using printed photogrammetric coded targets.

4. DATA PROCESSING

4.1 Range data processing

The 3D scanning pipeline (from multiple range maps to a complete 3D model) is nowadays a somehow standardized process. The main steps (registration of scans and merging) are available on most of the software associated to the acquisition devices, or alternatively covered by commercial or free tools (e.g. MeshLab, CloudCompare). No strong technical background is needed to acquire a *simple* object and generate a 3D model. However, when dealing with more complex objects, critical issues may arise not only in the planning and acquisition phases, but also during the processing one.

4.1.1 Issues in the registration step: The registration step has the aim of “reassembling” the shape of the object by aligning all the single acquisitions in a single reference space. This is usually obtained in a semi-automatic fashion, but the calculation of the initial alignment may need an input by the user. A first problem could be related to the required control over the coverage of the artwork surface. Most of the more recent 3D scanning devices offer software drivers that enable the user to align the acquisitions in real-time during digitization. But when the amount of scans becomes too big, it is not possible to complete the on-line alignment in a single session. Hence, a sufficient amount of overlapping and coverage of the object must be provided with a careful acquisition. In the case of the Sarcophagus, an almost complete coverage of the object was obtained, having a high redundant number of scans.

A second major issue is related to the overlap between single scans. In general, the alignment error tends to accumulate when the number of scans increases, and the best distribution of alignment error is obtained when the object is covered by a “closed” loop of scans. In the case of the Sarcophagus, the object was nearly “closed”, since its bottom part was not acquired. Hence, the highest degree of alignment error accumulation may have been concentrated in that part.

A third source of difficulties could be the necessity to combine different types of data, coming for example from different devices (or, in the case of the Sarcophagus, from different lenses). In this case, data with different resolution and amount of noise could be harder to align and control. In the case of the Sarcophagus, the highest amount of alignment error was found in the lower part of the Sarcophagus (Fig. 8a). In this area, scans coming from wide and medium lenses had to be combined, and unfortunately the amount of overlap was lower w.r.t. the rest of the object.

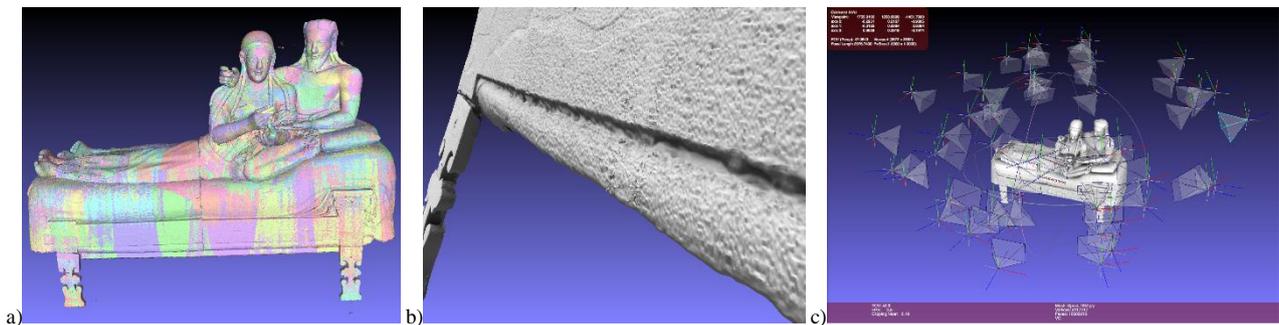


Figure 6. (a) Visualization of the aligned range maps. It can be noted that a minor redundancy and overlapping is present in the lower part of the Sarcophagus. (b) Artefacts generated by small misalignments of range maps. (c) Estimated camera positions for some of the images used for colour projection.

4.1.2 Issues in the merging phase: The merging step reconstructs a final 3D point cloud from the set of registered, highly redundant scans.

The number of scans could influence the choice of the reconstruction method and possibly lead to the necessity of simplifying or reducing the input data. For example, the methods based on implicit surfaces like the Poisson algorithm (Kazhdan et al., 2006) could be hard to handle when a very big dataset is available. On the other hand, volumetric methods based on marching cube can be applied, but could lead to the creation of geometric artefacts. The amount of noise in the original scans, especially when the original data come from different sources, may cause artefacts, or limit the size of the detail that can be reconstructed. In the case of the Sarcophagus, since the level of detail that had to be reconstructed was quite high, some portions of the surface coming from the wide lens of the scanner brought to the reconstruction of noisy surfaces (Fig. 6b), that needed a further improvement of registration quality, and also to the necessity of further cleaning of the original scans.

The whole geometric processing step (registration, merging, artifacts removal) took nearly two days of work. Half of this time required the active contribution of the user (registration, results checking, cleaning), while the rest was mainly unattended processing time.

The final model was produced using a volumetric merging method, implemented in MeshLab. This merging method uses a volumetric grid, the size of the cells (what is normally called the merging resolution) was set to 0.5mm. The merging resolution is normally chosen according to the needed final resolution, but it is also bound by the original data resolution and the residual alignment error (it is useless to choose a merging resolution lower than these values). In this case, the original sampling resolution was between 0.3-0.5mm, and the mean alignment error was estimated at 0.45mm. While the merging resolution was 0.5mm, given the noise in some of the scans, and the error residuals of alignment in some specific areas, we can estimate the accuracy of the model to be between 0.5 and 0.75mm. This value clearly varies across the surface: the two human figures have less noisy scans and lower alignment error, while most of the problems are located in the lower part of the Sarcophagus.

The final polygonal model was a triangulated mesh made of 61 million triangles. Starting from the “master model” we generated lower-resolution models (for interactive use, rendering and easier real-time manipulation) using geometric simplification.

4.1.3 Processing the colour data: While 3D scanners can produce accurate geometric data, most devices do not provide colour data, or in the case they do, the sampled colour is of low quality (usually due to the complexity of setting a proper illumination condition for both shape and colour acquisition). For this and other reasons, often the colour coverage comes from an ad-hoc photographic sampling. Unfortunately, this means that this photographic data needs an additional processing step: registration of the photographs over the 3D model and integration of those colour data over the triangulated surface.

The colour projection pipeline is a quite consolidated set of steps. The main ones are the image registration and the colour projection. While the latter can be obtained with automatic procedures, the former may need user intervention.

In the case of the Sarcophagus, a quite large number of images was acquired (306) and mapped using the procedure proposed in Corsini et al. (2013). The method makes use of Structure from Motion (SfM) to estimate the relative position of the cameras of the photographic set and a sparse reconstruction of the scene. The reconstruction is then used to retrieve the relative scale factor, and an initial registration of the images onto the geometry to-be-textured. The global registration is then refined to obtain an optimal solution (Fig. 8c). The colour projection and integration over the Sarcophagus mesh was applied using the Masked Photo Blending (Callieri et al., 2008) approach.

In the case of complex projects, the number of images and the amount of detail to be preserved may pose severe issues. The quality of the image registration and the possible artefacts produced by the colour projection must be carefully checked and corrected. Moreover, additional artefacts coming from occlusions may need the removal of some images from the dataset. While quite robust and scalable, the results of the colour projection pipeline are strongly related to the quality of the image dataset: this has an impact on the time needed for image shooting and in the amount of work to remove artefacts.

In the case of the Sarcophagus, the whole colour projection procedure took two days of work, roughly divided in: one day for SfM calculation, registration on the 3D model and refinement of registration; one day for colour projection and artefacts removal. The colour information was encoded in a colour-per-vertex fashion, since the high geometric resolution enabled to preserve colour detail.

4.2 Photogrammetric data processing

4.2.1 Image orientation and bundle adjustment: When processing large image datasets such as the one collected for the Sarcophagus, the processing procedure must be effective and, at the same time, reliable. The quality of the results can be controlled only if the user is able to access each step of the photogrammetric workflow and this is not always the case for automatic software packages. Furthermore, the nature of automatic image point observations extracted by SIFT-like operators, makes the reviewing of results very difficult as image observations are very different from what people would choose manually (crosses, corners, high contrast points, etc.). For these reasons, in the last years, many researchers have developed hybrid procedures, sometimes integrating commercially available packages with their own trusted procedures. The 3DOM-FBK team has been successfully using automatic commercial (Agisoft Photoscan, PIX4D, IwitNESS, Photomodeler) and research (GRAPHOS, APERO/MICMAC, VisualSFM) tools in the last years.

For the Sarcophagus the images were automatically oriented using Agisoft Photoscan, the extracted tie points, filtered and reduced in number using the procedure proposed by Nocerino et al. (2014), taking into account multiplicity of rays and their spatial distribution. The filtering procedure was necessary to keep only well distributed and reliable image correspondences and to reduce the computational time when using bundle adjustment software. Tie points were then imported in Australis and a self-calibrating bundle adjustment was run to obtain main statistical values and inter-correlations between additional and exterior orientation parameters (RMS of image residuals resulted 0.43 pixel - with a max of 2.17 pixel - whereas the average intersecting angle resulted 61.1°).

Figure 7 shows the filtered 3D tie points are coloured according to: (a) the precision vector length, (b) the number of images in which each tie point is visible (i.e. multiplicity).

All the images were processed in a unique self-calibrating bundle adjustment, providing the possibility of using the scale bar in the different positions for accuracy assessment. The analysis of residuals on the scale bars allowed to find out that one of the scale bar moved during the acquisition probably because of the soft covering on the Sarcophagus stage. Removing from the calculus the unstable scale bar, a mean absolute error less of some 0.08 mm was obtained on the remaining reference lengths.

To check the mechanical stability of the camera-lens system, a further processing was performed using as interior orientation parameters the values computed from the camera calibration processing. The mean absolute error on the scale bars did not significantly vary from the self-calibrating adjustment, meaning that the photographic system remained stable during the image acquisition.

4.2.2 Dense image matching: In general, image matching can be defined as the establishment of correspondences between two or more images (Schenk, 1999). Specifically, dense image matching refers to the procedure of generating 3D point clouds of the acquired object at a density comparable to the resolution of the images. In the last years thanks to improvements in hardware and algorithms, image-based modelling has become a valid alternative to range sensors in terms of attainable accuracy and resolution (Remondino et al, 2014). However, dense image matching requires highly demanding computational efforts and, therefore, often employs advanced technique, such as parallel computing and GPU/FPGA implementation. As easily understandable, the quality of the results is highly affected by the quality of both images and camera network. Despite the developments of various matching strategies, issues still exist: (i) textureless, in-shadows areas or images with low signal-to-noise ratio which lead to noisy point clouds or no reconstruction; (ii) geometrical edges which may not be preserved thus resulting in smoothing effects; (iii) mixed pixels on object boundaries which lead to erroneous points (Fig. 8a); (iv) lack of quality indicators which are in most of the cases not provided; (v) absence of reliable automatic filtering procedures for noisy point clouds - consequently the cleaning operations may be time consuming and may require heavy computational power.

For the Sarcophagus, the dense image matching phase employed all available images simultaneously in order to improve the dense reconstruction and allow filtering of mixed points thanks to the intersection of multiple matching rays in object space (Fig. 8b). The inter-point spacing, or lateral resolution, of the final 3D point cloud (Fig. 8c) was better than 0.3 mm.

4.2.3 Mesh generation and editing: The generation of surfaces from point clouds is a fundamental step in the digitization and modelling pipeline. It is hard to provide a comprehensive classification of reconstruction algorithms (Remondino, 2003). However, generally speaking, surface reconstruction methods try to *wrap* a surface over the points keeping their original coordinates (Ramos and Sadri, 2007). For the mesh generation of the Sarcophagus, the Poisson algorithm implemented in CloudCompare provided the most effective meshing results. A high resolution (≈ 0.5 mm) mesh model composed of more than 50 million triangles was generated.

Triangulated meshes usually need some editing operations to fix topological errors (such as self-intersecting triangles, holes, etc.). Such editing procedures are still manual to a great extent. In Rodríguez-González et al. (2015) a simplification and optimization workflow for generating lightweight models that keeps the maximum amount of significant information is proposed. The same procedure was applied to the high resolution model of the Sarcophagus in order to get a lower resolution mesh (≈ 2 mm) of 250,000 triangles, more suitable for visualization purposes.

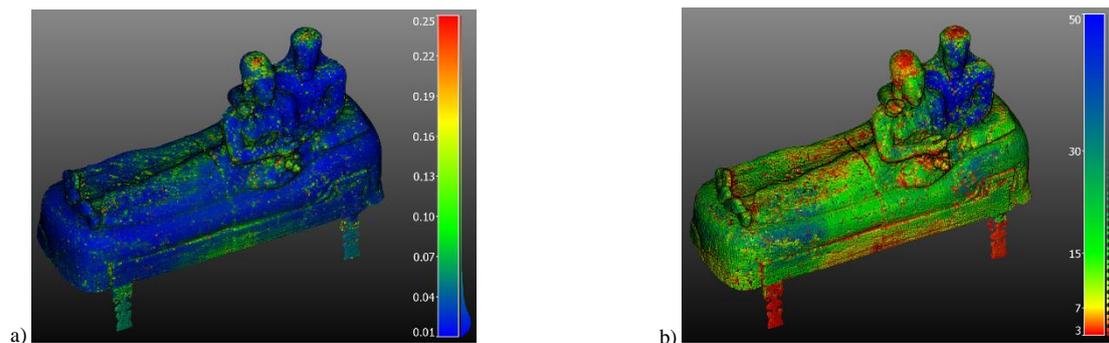


Figure 7. Colour-coded visualization of the filtered 3D tie points according to: a) precision vector length (mean = 0.057 mm, max = 1.265 mm); b) multiplicity, i.e. number of images where a 3D point is visible (mean = 14, max = 98, min = 3).



Figure 8. a) Noisy points generated from mixed pixels on boundaries. b) The erroneous points are reduced when all the images are used together. c) Particular of the photogrammetric textured 3D model.

4.2.4 Texture mapping: Texture mapping is a graphic design process in which a 2D surface, called a texture map, is "wrapped around" a 3D object. Texture mapping is the electronic equivalent of applying wallpaper, paint, or veneer to a real object (<http://whatis.techtarget.com/definition/texture-mapping>). In photogrammetry, from a geometrical point of view, the process is straightforward: the collinearity principle is adopted to map colour information from images to the mesh. However, a texture mapping algorithm has to consider and solve the following issues: (i) occlusions must be checked through visibility analysis in order to identify the triangles of the mesh actually seen by the images and avoid erroneous projections; (ii) when multiple images *look* at the same triangles, the ones with most perpendicular optical axis must be selected; (iii) colour blending, necessary to guarantee an homogenous texture, especially when lighting conditions had changed during the image acquisition; (iv) colour balancing may be required to provide faithful colour appearance to the mesh model. An advanced algorithm should also consider lens vignetting, DOF issues, shadows and lens aberrations across the image format. Figure 8c shows a particular of the final texturized 3D model obtained through a procedure that takes into account the above mentioned requirements.

previous sections, describing the planning and acquisition steps of both technologies, it is clear that, beside the "ideal" and "standard" procedures, the work on the field heavily rely on a series of precautions, safeguards and redundancy strategies, mostly deriving from experience and on the knowledge of the weak points and recurring problems of both technologies. Similarly, some parts of the processing phase are tweaked to accommodate the specific characteristics of the dataset and are heavily influenced by what has emerged, during digitization, as critical areas of the target object.

Considering the data acquisition and processing efforts (Table 1), photogrammetry resulted slightly longer, probably due to the employed very high-resolution images (24 Mpx).

	<i>Laser scanning</i>	<i>Photogrammetry</i>
<i>Acquisition</i>	ca 7 hours	ca 13 hours
<i>Numb. data</i>	351 scans, 306 images	520 images
<i>Processing</i>	ca 4 days	ca 5 days

Table 1: A summary of the efforts for both techniques.

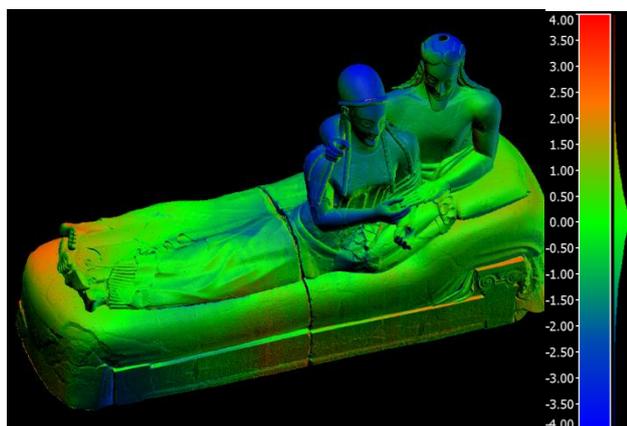


Figure 9. Euclidean distances between the two 3D models digitized using photogrammetry and laser scanning. While on most of the artefacts the 2 approaches delivered very similar results, some local deformations are present in the upper part of the two figures.

For complex CH assets such as the Sarcophagus, local deformations are likely to occur and can be expected despite rigorous approaches are followed during surveying and processing stages. Such a behaviour has been highlighted also for the Sarcophagus (Fig. 9) where some local surface deformations up to 3-4 mm are visible – although the mean value of the distance between the two 3D models is 0.1 mm (std=0.9 mm). This phenomenon is still under investigation and will be analysed more deeply in a successive work. Such differences may be originated from different causes such as error accumulation in bundle block adjustment or range map alignment, mechanical instability of the scanner in rolled positions or residual systematic errors not properly modelled in the calibration.

ACKNOWLEDGMENTS

Authors are thankful to Soprintendenza per i Beni Archeologici dell'Etruria Meridionale and CINECA. A special thank also to Manfrotto, a Vitec Group company, for providing the photographic materials (lights, clamps, tripods, etc.) used in the photogrammetric survey.

5. COMPARISONS AND DISCUSSION

When photogrammetry and 3D scanning are normally described in books and in papers, they may appear as technologies with a completely fixed protocol, where the procedures are always perfectly standardized. This may be considered true only for the more *simple* digitization campaigns. However, by looking at the

REFERENCES

Alsadik, B., Remondino, F., Menna, F., Gerke, M., Vosselman, G., 2013. Robust extraction of image correspondences exploiting the image scene geometry and approximate camera orientation. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 40, pp.1-7.

- Callieri, M., Cignoni, P., Corsini, M., Scopigno, R., 2008. Masked Photo Blending: mapping dense photographic dataset on high-resolution 3D models. *Computer & Graphics*, Vol. 32(4), pp. 464-473.
- Callieri, M., Chica, A., Dellepiane, M., Besora, I., Corsini, M., Moyés, J., Ranzuglia, G., Scopigno, R., Brunet, P., 2011. Multiscale acquisition and presentation of very large artifacts: The case of Portalada. *ACM Journal on Computing and Cultural Heritage (JOCCH)*, Vol. 4(4).
- Cignoni, P., Rocchini, C., Montani, C., Scopigno, R., 2003. External Memory Management and Simplification of Huge Meshes. *IEEE Transaction on Visualization and Computer Graphics*, Vol.9 (4), pp. 525-537.
- Corsini, M., Dellepiane, M., Ganovelli, F., Gherardi, R., Fusiello, A., Scopigno, R., 2013. Fully Automatic Registration of Image Sets on Approximate Geometry. *International Journal of Computer Vision*, Vol. 102(1-3), pp. 91-111
- Fraser, C.S., 1996. Network design. In “*Close-range Photogrammetry and Machine Vision*”, Atkinson (Ed.), Whittles Publishing, UK, pp. 256-282.
- Kazhdan, M., Bolitho, M., Hoppe, H., 2006. Poisson surface reconstruction. In *Proc. 4th Eurographics Symposium on Geometry Processing*.
- Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., 2000. The digital Michelangelo project: 3D scanning of large statues. In *Proc. 27th Computer Graphics and Interactive Techniques*, pp. 131-144, ACM Press/Addison-Wesley Publishing Co.
- Nocerino, E., Menna, F., Remondino, F., 2014. Accuracy of typical photogrammetric networks in cultural heritage 3D modeling projects. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(5), p.465.
- Ramos, E.A. and Sadri, B., 2007. Geometric and topological guarantees for the wrap reconstruction algorithm. In *Proc. 18th annual ACM-SIAM Symposium on discrete algorithms*, pp. 1086-1095. Society for Industrial and Applied Mathematics.
- Ray, S.F., 2002. *Applied photographic optics: Lenses and optical systems for photography, film, video, electronic and digital imaging*. Focal Press.
- Remondino, F., Menna, F., Koutsoudis, A., Chamzas, C., El-Hakim, S., 2013. Design and implement a reality-based 3D digitisation and modelling project. *Proc. IEEE Conference "Digital Heritage 2013"*, Vol. 1, pp. 137-144.
- Remondino, F., Fraser, C., 2006. Digital camera calibration methods: considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 36(5), pp. 266-272.
- Remondino, F., 2003. From point cloud to surface: the modeling and visualization problem. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34(5), p.W10.
- Remondino, F., Spera, M.G., Nocerino, E., Menna, F. and Nex, F., 2014. State of the art in high density image matching. *The Photogrammetric Record*, 29(146), pp.144-166.
- Rodríguez-Gonzálvez, P., Nocerino, E., Menna, F., Minto, S. and Remondino, F., 2015. 3D surveying & modeling of underground passages in WWI fortifications. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(5), p.17.
- Russo, A., Cosentino, R., De Lucia, M.A., Guidazzoli, A., Cohen, G.B., Liguori, M.C. and Fishnaller, F., 2013. Etruscan cultural heritage: The Sarcophagus of the Spouses project: Virtual storytelling embedded in sensory audiovisual environments. In *Proc. IEEE Conference "Digital Heritage 2013"*, Vol. 1, pp. 441-441.
- Schenk, T., 1999. *Digital Photogrammetry: Volume I*. TerraScience, Laurelville, Ohio, USA. 421 pages.