

COMPUTER VISION TOOLS FOR 3D MODELLING IN ARCHAEOLOGY

M. Lo Brutto*, P. Meli

Dept. of Civil, Environmental, Aerospace and Materials Engineering
University of Palermo, Viale delle Scienze, 90128 Palermo, Italy
(mauro.lobrutto, paola.meli)@unipa.it

KEY WORDS: Computer Vision, 3D Web Service, Archaeology, 3D Modelling, Laser scanning, Survey

ABSTRACT:

In archaeological Cultural Heritage study 3D modelling has become a very useful process to obtain indispensable data for documentation and visualization. Nowadays the continuous request to achieve photorealistic 3D models has led to testing different techniques and methodologies to speed up both data acquisition and the data processing phase. There are many examples of surveys conducted with the use of range-based and image-based techniques, but, in the last few years, the scientific research has been increasingly moving towards automatic procedures using Computer Vision approach to reduce time during data processing. Computer Vision approach offers a great opportunity for archaeological survey since it can be very easily used by existing Computer Vision interfaces such as 3D web services and open source or low cost software. The aim of this work is to evaluate the performance offered by Computer Vision interfaces for 3D survey of archaeological ruins using some 3D web-service tools and a low cost software like PhotoScan package. Some tests have been performed to analyze the geometric accuracy of 3D models obtained by 3D web-service tools and PhotoScan package through the comparison with a 3D model achieved by laser scanning survey.

1. INTRODUCTION

The study of archaeological Cultural Heritage has a valid support in new technologies offered by geomatics to create highly detailed and accurate 2D/3D products. Indeed in this field the construction of detailed 3D models for different types of representation and documentation is increasingly required in order to carry out analysis for restoration, for historical studies or simply for visualization. For these reasons, the research is increasingly directing towards the study of methodologies able to enhance the performance of the geomatics techniques and to reduce the time both in survey and in processing phase.

In Cultural Heritage documentation, and particularly in archaeology documentation, the survey techniques should have some properties such as accuracy, low cost, portability and rapidity of data acquisition (Remondino & Rizzi, 2010). Even if the range-based techniques are in general more accurate, the image-based techniques are more convenient and practical (Remondino & El-Hakim, 2006; Bitelli et al.; 2007, Barazzetti et al., 2011). Furthermore, the image-based techniques, through the use of algorithms derived from Computer Vision (CV), known as "Structure from Motion", are able to automatically perform the whole pipeline reducing time both of images orientation and 3D reconstruction (Vergauwen & Van Gool, 2006; Barazzetti et al., 2011; Doneus et al. 2011;). The "Structure from Motion" approach allows to orient a very huge numbers of images without any knowledge of the camera parameters and network geometry (Barazzetti et al., 2010). The images orientation is performed automatically identifying the common feature points through appropriate interest operators. The SIFT (Scale Invariant Feature Transform) (Lowe, 2004) is one of more interest operator used in CV because it allows to find correspondences between images taken from different positions, with different scales and different illuminations. The camera parameters can be estimated during the matching phase (self-calibration). In the "Structure from Motion" approach all consecutive images are coupled and the subsequent image is

matched with the previous pairs. Since all the images are matched together the whole epipolar geometry is reconstructed and thus the projective scene geometry. The results of this process are the camera parameters, the images orientation and a sparse point cloud of the object. These points are not useful to obtain a satisfactory reconstruction of the model but they constitute the basis for the subsequent phases of image matching, such as the dense stereo matching techniques, which allow to create a very detailed point cloud.

The "Structure from Motion" has been used in many applications like reverse engineering (Menna & Troisi, 2010) or UAV applications (Nietzel & Klonowski, 2011). Some interesting experiences have been done combining spherical photogrammetry and "Structure from Motion" tools (D'Annibale et al., 2011).

The "Structure from Motion" approach is used in many open source software (Bundler, CMVS/PMVS2), in some photogrammetric commercial packages (PhotoModeler 2012 by EOS Systems Inc., PhotoScan by AgiSoft LLC) and in many 3D web services (Autodesk 123D Catch, ARC 3D Webservice, Photosynth from Microsoft, Hypr3D). The 3D web services are web tools for remote 3D reconstruction; they allow the reconstruction of 3D models from images in a very short time and without any specific knowledge. These services are available online and are completely freeware. They are created for inexperienced users and mainly for visualization but they could be very useful especially in archeology field, where generally the campaigns have restricted budget.

Commonly, CV techniques aim more on process automation rather than on precision and accuracy (Barazzetti et al., 2010). This condition is disadvantageous when the purpose of the 3D survey is also a model with a good metric accuracy. Nevertheless promising results have been achieved for the survey of small objects using ARC3D (De Balestrini & Guerra, 2010), and of a prehistoric site using Autodesk 123Dcatch (Chandler & Fryer, 2011).

* Corresponding author.

The aim of this work is to evaluate the performance offered by the "Structure from Motion" approach for 3D survey of archeological ruins using existing CV interfaces such as some 3D web-services and PhotoScan package. Some tests have been performed to analyze the geometric accuracy for the creation of 3D models through the comparison with a 3D model obtained by laser scanning data. The tests have been direct to obtain a 3D model of the ruins of the archaeological site of Solunto; in particular, in this first stage we have only surveyed the ruins of the ancient Greek theatre. The ancient city of Solunto is situated on the north coast of Sicily, near Palermo, on the south-east side of Monte Catalfano. It was founded by Phoenicians in the VI century A.C.; later it became a greek colony and finally a roman colony. The city had a Hellenistic urban plan of "Hippodameic" type, where the three main centres, religious, administrative-political and commercial, were included in the layout of equal-sized blocks of buildings delimited by straight, often paved, roads with perpendicular intersections.

The theatre is located above the "agora" and it originally had a diameter of about 45 m and 21 orders of seats; the theatre could receive up to one thousand people (Figure 1). The orchestra has two overlapping floors dating in two subsequent phases: the former is probably of the VI century B.C., the latter of the Hellenistic period. Today only a little portion of tiers remains.



Figure 1. The theatre of Solunto

The first part of the article describes the laser scanner and the images acquisition; the second part reports the data processing. The 3D models were carried out using different 3D web-services (Photosynth, ARC3D, Hypr3D, Autodesk 123Dcatch) and a low cost commercial software (PhotoScan by AgiSoft LLC). In the last part of the article some comparisons have been performed to evaluate the quality and metric accuracy of the 3D models generated with image-based approach.

2. DATA ACQUISITION

The 3D survey of theatre of Solunto was conducted using topographic, laser scanning and image-based techniques. The topographic survey, executed by using a Leica 1105 total station from a single station point located in front of the theatre, was used to measure the coordinates of some laser scanning and photogrammetric targets. In this way we can refer all 3D data to the same reference system. The laser scanning survey enabled to cover the area once occupied by the "cavea" of the ancient theatre, while the photogrammetric survey covered only a little portion of the tiers of seating.

2.1 Laser scanning acquisition

The laser scanner survey was conducted using a FARO Focus^{3D} S120. This laser scanner uses phase shift technology and has a measuring range from 0.6 to 120 m (Figure 2).



Figure 2. Laser scanner survey

Five scans were executed from five station points: three under the "cavea" and two from the top of the "cavea". Eight checkerboard targets were placed in the area of the theatre in order to merge all the scans in the same local reference system. The scans resolution was set with a step of 5 mm at a distance of 10 m. Finally the portion of the model representing the theatre seats was used as a reference model for the tests.

2.2 Images acquisition

The images acquisition was conducted in two different steps using a Nikon D80 digital camera equipped with fixed focal length lens of 28 mm with a sensor size of 23,6 mm x 15,8 mm (3872 x 2592 pixels) and a pixel size of 6.1 μm. The first step was performed during the laser scanning data acquisition and the topographic survey. Five convergent images were taken from a distance of about 5 m. These images weren't suitable for a 3D reconstruction, but they were used only to scale and to rototraslate the 3D web service models.

The second one was performed with a telescopic pole having a bayonet cap, where the digital camera was applied on top. The telescope pole is 2 m long when fully extended. The images were taken putting the pole at different heights: some of them placing the pole on the ground, others raising it to height more than 2 m high. The images were both convergent and stereoscopic and they were taken tilting the camera to obtain a set of images providing a good coverage of the object. Three stereoscopic strips were taken from three different heights: the first one from a height of about 2 m, the second one from a height of about 2,25 m, the third one from a height of approximately 2,50 m. A total of 25 stereoscopic images were taken from an average distance of about 4 m. The convergent images were taken from a height of 2,00 m and a distance of 6 m; totally 17 convergent images were taken (Figure 3).

The images taken during the second step were used to obtain the 3D reconstruction. They were captured one week after the first step; these images didn't have the photogrammetric targets (they were removed) useful for images orientation.

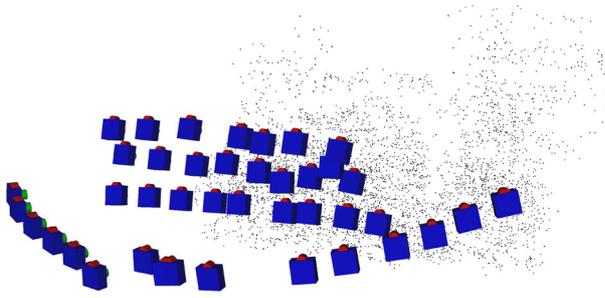


Figure 3. Camera network of images used for 3D reconstruction

3. DATA PROCESSING

3.1 Laser scanning data processing

The laser scanner data processing was carried out with the software Faro Scene and Geomagic Studio. In the first step every scan was registered and merged with Faro Scene software. Before the registration procedures a first editing was conducted by removing the external parts and those portions with presence of vegetation. The scans were aligned and registered to the global reference system using the eight targets obtaining a standard deviation of 4 mm. The point clouds were exported in ASCII and were processed using Geomagic. Some filtering was applied to remove noise, redundant points and outliers. Then all the scans were merged together. Finally a polygonal model (“mesh”) of the theatre seats was created.

3.2 The Computer Vision tools

The image data processing was executed with the 3D web-services Photosynth, ARC3D, Hypr3D, Autodesk 123Dcatch and with the low cost commercial software PhotoScan.

Photosynth is a Microsoft web service which is accessible through a Windows Live account at the Photosynth website. It requires the installation of an application to allow the uploading of the images to the server. Photosynth enables to create two kinds of 3D products: panoramas and synths. The former stitches a set of images together taken from the same point to create a seamlessly panoramic picture. The latter creates a view that allows to browse from photo to another photo by using a set of overlapping images. This process was created especially for “Photo Tourism” (Snavely et al., 2007), but it is also able to create a point cloud of the object. The camera calibration parameters, the images exterior orientation parameters and the point cloud can be exported through SynthExport applications (<http://synthexport.codeplex.com>) in different formats.

ARC3D (Automatic Reconstruction Conduit) is a EPOCH network tool for creating 3D models (point clouds and polygonal model) out of a set of images. The user must install a software to upload the images to the server. If the reconstruction has been successful the 3D model can be downloaded in different formats. The model can be displayed through the use of MeshLab, an open source software used for viewing and editing point clouds and 3D triangulated models. To obtain good results the images should be convergent with angle less than 10 degrees and large overlap.

Hypr3D works directly online with images or videos. The first step is the choice of file format to upload (images or videos). The second is the upload of the files, at least five images, that

are useful for the creation of the 3D model. The process computes the camera parameters, produces a point cloud, a wireframe model and a texturing high-resolution model. The model can be downloaded in different formats and resolutions.

Autodesk 123Dcatch is a web service which is managed through a software downloaded from www.123dapp.com/catch. With this software it is possible to upload the images, to define the camera focal length and to modify the photo-scene after the first process. An important difference compared to the other web services is the equipment of some tools that allow to modify or to improve the result, e.g. the manual collimation of points to orient the images or to scale the model. Furthermore it is possible to set the origin and the axes of a global reference system. The final product is a triangulated surface which can be achieved in three different resolutions: mobile (fast medium resolution mesh), standard (high resolution texture mesh) and maximum (very high density mesh).

Recently two kinds of apps are available: the web app that allows to compute the 3D scene directly on line and the iPad app that enables to create and to view the 3D scene on iPads.

Agisoft PhotoScan is a low cost image-based package aimed to obtain high quality 3D model. The software is based on multi-view 3D reconstruction technology and can operate with calibrated and un-calibrated images in both controlled and uncontrolled conditions. The workflow includes fully automatic images orientation and 3D model reconstruction; the model can be exported for editing in external software. All the processes can be performed with different levels of accuracy and many parameters can be set to improve the final result. For our work we used the PhotoScan professional edition installed on a 32-bit system. This condition could limit the software performance especially when the product must be a high quality 3D model.

3.2.1 Images data processing

Microsoft Photosynth: The model created with Photosynth was performed using all stereoscopic and convergent images (42 images). The 3D web-service was able to orient all the images in a very short time. This is a big advantage when it is necessary to orient a large number of images. However the produced point cloud was very noisy and sparse, sometimes inadequate to obtain a detailed 3D model. Moreover a limitation of this web service is the absence of tools that allow to create a polygonal model (mesh). Overall about 120.000 points have been calculated. The points were imported in Geomagic to produce a polygonal model. The poor point cloud has allowed to obtain a 3D model composed by a number of triangles lower than that obtained by the other interfaces. The mesh was poorly defined with the presence of holes mainly in correspondence of the upper surfaces of the higher seats and in the lateral surfaces of the seats. An editing phase was necessary and a 3D model with about 230,000 triangles was achieved (Figure 4).

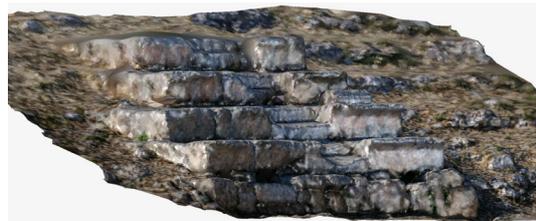


Figure 4. Photosynth 3D model

Arc3D: The experience made with Arc3D has highlighted some problems during the images processing. To obtain the 3D model several tests have been executed using different images block with different levels of re-sampling. The first test was performed using all the images without applying any images re-sampling. The 3D web-service was able to upload all the images, but the obtained model was incomplete. Therefore other tests were carried out using all the images and applying a re-sampling up to 50% of the original images resolution without obtaining satisfactory results. Other tests were performed with only the 25 stereoscopic images and only the 17 convergent images. The best result, a high resolution 3D model in regard to both the texture and the level of detail, was obtained using only the 17 convergent images without performing any resampling. The automatic process was able to reconstruct completely the surfaces of the theatre seats. There were holes only on the upper surfaces of the higher seats. The model was imported in Geomagic to perform some editing operation in order to obtain a complete 3D model composed of 390,000 triangles (Figure 5).



Figure 5. ARC3D 3D model

This experience has allowed to state that often the process can fail with this web service, when the used images don't meet the requirements for the processing. Furthermore it is not so simple to understand the problem of the process failure, so to obtain good results it is necessary to execute many tests. However when the 3D reconstruction has been successful, it is possible to obtain high quality 3D models in a short time.

Hyper 3D: Also with Hyper 3D some tests were performed with different set of photos to evaluate what kind of images is better to obtain a complete and high resolution 3D model. Three tests were carried out: the first one using all the 42 images, the second one using only the 25 stereoscopic images, and the last one using only the 17 convergent images. Hyper 3D was able to orient and to obtain a good 3D model with all the used datasets, but the 3D model obtained by the 17 convergent images was better than the others and it was used for the subsequent comparisons. A point cloud of about 150,000 points was obtained. As Photosynth and Arc 3D this web service doesn't have any tools to modify or to improve the obtained 3D model so the point cloud was imported in Geomagic to create a polygonal model. Some holes were present in less visible portions from the images. After an editing phase a mesh with about 220.000 triangles was obtained (Figure 6).



Figure 6. Hypr 3D model

Autodesk 123Dcatch: All images have been used with this web service and their orientation has been successful, but the obtained 3D model showed some imperfections, so to refine the orientation phase and to improve the 3D reconstruction it was necessary to measure manually some points. Indeed, after the first processing, it is possible to choose some points directly on the images to help the process orientation. This possibility is a great advantage, as it makes the user interaction possible to aid the 3D reconstruction. Another advantage is the opportunity to set the resolution of the mesh according to the aim of the survey. For this work the 3D model was computed with maximum resolution (very high density mesh). Any editing phase was performed as the 3D reconstruction appeared satisfactory and there were not holes. A polygonal 3D model was created with about 800,000 triangles (Figure 7). Finally a big disadvantage is the absence of tools that allow to check the correct orientation of the images, as will be seen below, this can create problems in geometric accuracy of the final model.



Figure 7. Autodesk 123Dcatch 3D model

Photoscan: The data processing carried out with Photoscan was executed using all the images (42 images). The images were automatically oriented performing a first orientation with a low level of accuracy and afterwards re-computing the orientations with a high level of accuracy. This packages is different from the previous tools, because it is a commercial software that doesn't work on line. Moreover it is possible to set the accuracy of every phase of the whole pipeline, from the images orientation to the 3D reconstruction. Several editing commands to modify or to improve the obtained 3D model are available, such as mesh decimation, removal of detached triangles, closing of holes, etc. Photoscan is able to produce a 3D model with very high geometry resolution, but, to obtain this, it is necessary to have a high-performance workstation. Consequently the resolution of the obtained 3D model has been limited by the used hardware. Therefore in our case it was possible to compute a 3D model with a medium target quality and imposing a maximum limit of 400.000 triangles (Figure 8). The obtained level of detail was considered sufficient for this work.



Figure 8. Photoscan 3D model

4. METRIC ACCURACY EVALUATION

In order to evaluate the metric accuracy of the 3D models a comparison was performed using the laser scanner 3D model as reference.

The workflow was:

- to scale the web service 3D models using a reference distance;
- to align the web service 3D models in the global reference system;
- to compare laser scanner and web service 3D models.

To scale the web service 3D models some reference distances were measured from a photogrammetric project done with Photomodeler Scanner using the first set of images, containing the photogrammetric targets surveyed with the total station. Photosynth, Arc3d and Hypr3D models were scaled adding directly on the 3D models a distance between two points; Autodesk 123D catch model was scaled adding a distance directly on the images. The alignment of the web service 3D models in the same reference system of the laser scanner 3D model was performed with Geomagic through the ICP (Interactive Closet Point) registration algorithm.

The Photoscan 3D model was scaled and georeferenced in the global reference system identifying four markers on the images. The markers coordinates were computed through the photogrammetric project of PhotoModeler Scanner.

The third phase was conducted with the software Geomagic to perform a 3D comparison between the surfaces of web service models and laser scanner model.

The standard deviations resulting from the comparison with the models obtained with Arc3D and Hypr3D are 0,012 m while those resulting from the comparison with the models created with Photosynth and Photoscan are 0,014 m. The worst result was achieved with Autodesk 123Dcatch; the standard deviation resulting from this comparison is 0,026 m.

A 3D map, that shows the distribution of deviations between the different models through a color scale, was created for each model; the deviations represent the shortest distance from the reference model in 3D. The distribution of deviations is not uniform for all the models. The 3D maps of Photosynth, Arc3D, Hypr3D, and Photoscan show that the maximum deviations are $\pm 0,06$ m and they are situated in critical areas such as the interstices among the stones of the seatings, where the data appears noisy, and the border areas, that are not visible on the images (Figures 9 - 12).

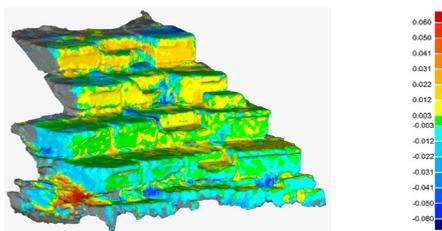


Figure 9. The 3D map of the comparison with Photosynth

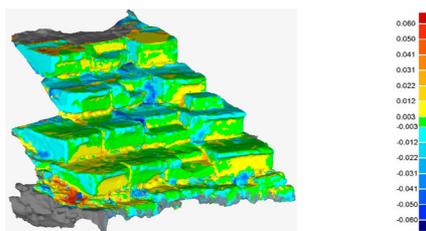


Figure 10. The 3D map of the comparison with Arc 3D

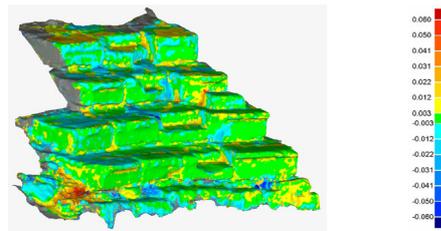


Figure 11. The 3D map of the comparison with Hypr3D

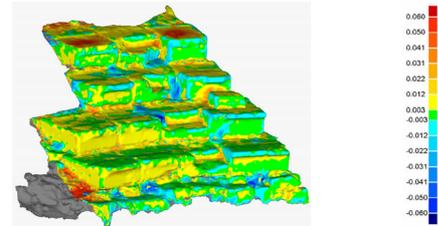


Figure 12. The 3D map of the comparison with Photoscan

Furthermore grey areas represent the portions where the two model don't match. These areas are mainly situated where the models had some holes that were filled manually.

The model showing the worst result was achieved with 123Dcatch. There is an odd distribution between the positive and negative values of the deviations both among seats of the same row and among seats of different row (Figure 13). The front surfaces of the seats in the two lower orders have positive deviations in the outer parts and negative deviations in the inner parts. This could indicate a problem on the correct geometry of the 3D model. Furthermore in the two upper tiers the deviations are for the most part negative.

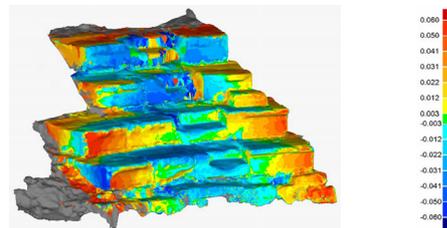


Figure 13. The 3D map of the comparison with Autodesk 123Dcatch

5. CONCLUSION

The experience described makes it possible to offer some considerations about advantages and disadvantages on the use of some CV interfaces for 3D survey in archaeology. The 3D web-services and PhotoScan package used for this work are promising tools to use in 3D survey of cultural heritage when there are short time and a limited budget.

Indeed the research showed that these instruments create 3D models with high resolution in a easy way and at low cost, but the automatic procedures are difficult to control due to the lack of tools able to check the 3D reconstruction. Sometimes the process can fail when the used images don't meet the requirements for the processing, for example in the experience executed with Arc3D where it was necessary to perform several tests to obtain an acceptable result. Furthermore the absence of editing commands that allow to scale and to georeference the 3D model directly in the web services is a big disadvantage. Only two among the used CV interfaces allow some editing

operations: Autodesk 123Dcatch and Photoscan. The first one has some tools that allow to perform some limited editing operations, while the second one has several editing commands. Moreover another disadvantage is the absence of parameters to verify the correct images orientation.

The standard deviations obtained from the comparisons show a centimeter accuracy that is too low for high accuracy applications. The 3D maps resulting from the metric evaluation show an uneven distribution of the deviations. This could be due to problems of scale and/or of an incorrect alignment between the web-service 3D models and the laser scanner 3D model. Some problems noted in the 3D model created with 123Dcatch could be due to the incorrect images orientation.

However these applications are useful but don't always yield a correct 3D reconstruction. The high photorealistic quality allows to use web service 3D models in archaeology mainly for visualization, documentation and multi-temporal studies. Surely these tools cannot replace laser scanning or photogrammetric techniques when it is necessary to have 3D models with high level of accuracy. Further tests should be done using different datasets to evaluate the real performance of these applications and to study some simple procedures to check the metric accuracy.

REFERENCES

References from Journals:

Barazzetti, L., Scaioni, M., Remondino, F., 2010. Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation, *The Photogrammetric Record*, 25(132), pp 356–381.

Barazzetti, L., Binda, L., Scaioni, M., Taranto, P., 2011. Photogrammetric survey of complex geometries with low-cost software: Application to the “G1” temple in Myson, Vietnam, *Journal of Cultural Heritage*, 12(3), pp. 253-262.

Lowe D., 2004. Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, 60(2), pp. 91-110.

Remondino, F., El-Hakim, S., 2006. Image-based 3D modelling: a review, *The Photogrammetric Record*, 21(115), pp. 269-291.

Remondino, F., Rizzi, A., 2010. Reality-based 3D documentation of natural and cultural heritages sites – techniques, problems, and examples, *Applied Geomatics*, 2(3), pp. 85-100.

Snavey, N., Seitz, S. M., Szeliski, R., 2007. Modeling the World from Internet Photo Collection, *International Journal of Computer Vision*, 80(2), pp 189-210.

Vergauwen, M., Van Gool, L., 2006. Web-based 3D reconstruction service, *Machine Vision and Application*, 17, pp. 411-426.

References from Proceedings and Books:

Bitelli, G. Girelli, V.A., Remondino, F., Vittuari, L., 2007. *The potential of 3D techniques for Cultural Heritage object documentation*. Proc. of Videometrics IX - SPIE-IS&T Electronic Imaging, Vol. 6491.

D'Annibale, E., Piermattei, L., Fangi, G., 2011. *Spherical photogrammetry as emergency photogrammetry*. XXIIIrd International CIPA Symposium, Prague, 12-16 September 2011.

Doneus M., Verhoeven, G., Fera, M. Briese, Ch., Kucera, M., Neubauer, W., 2011. *From deposit to point cloud – a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations*. XXIIIrd International CIPA Symposium, Prague, 12-16 September 2011.

References from Other Literature:

Fratu de Balestrini, E., Guerra, F., 2010. New instruments for survey: on line softwares for 3D reconstruction from images. In: *The International Archives of the photogrammetry, Remote sensing and Spatial Information Science*, Vol. XXXVIII-5/W16.

Menna, F., Troisi, S., 2010. Low cost reverse engineering techniques for 3D modeling of propellers. In: *The International Archives of the photogrammetry, Remote sensing and Spatial Information Science*, vol XXXVIII, Part 5, pp. 452-457.

Nietzel, F., Klonowski, J., 2011. Mobile 3D mapping with a low-cost system. In: *The international Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII-1/C22.

References from websites:

Chandler J., Fryer J., 2011. Accuracy of Autodesk 123D Catch? Aboriginal cave re-measurement using digital photogrammetry, <http://www-staff.lboro.ac.uk/~cvjhc/otherfiles/accuracy%20of%20123dcatch.htm>.

ACKNOWLEDGEMENT

The work presented in this paper was developed within the research project “Advanced methodologies for survey, documentation and visualization of Cultural Heritage”, supported by MURST (Italian Ministry of University and Research), as a Project of Relevant National Interest (PRIN2008). Local coordinator: Prof. Mauro Lo Brutto (University of Palermo). National Coordinator: Prof. Raffaella Brumana (Politecnico di Milano).

The authors would like to thank Dr Adriana Fresina, Director of the archaeological site of Solunto for allowing the access to the site, the Geotop s.r.l. for the availability of FARO laser scanner and Arch. Fabio Di Salvo for helping during the survey phase.