FROM PIXEL TO MESH – ACCURATE AND STRAIGHTFORWARD 3D DOCUMENTATION OF CULTURAL HERITAGE FROM THE CRES/LOŠINJ ARCHIPELAGO

UDK/UDC 902:004.89(497.572)
Pregledni članak
Primljeno: 29. 10. 2013.

Geert Verhoeven
VIAS – Vienna Institute
for Archaeological Science
University of Vienna
Franz-Klein-Gasse 1
A – 1190 Vienna
geert.verhoeven@archpro.lbg.ac.at

Michael Doneus
Department for Prehistoric
and Historical Archaeology
University of Vienna
Franz-Klein-Gasse 1
A – 1190 Vienna
michael.doneus@univie.ac.at

Nives Doneus
Ludwig Boltzmann Institute
for Archaeological Prospection
and Virtual Archaeology
Franz-Klein-Gasse 1
A – 1190 Vienna
nives.doneus@archpro.lbg.ac.at

Seta Štuhec
University of Ljubljana
Faculty of Arts
Aškerčeva 2
SLO – 1000 Ljubljana
seta.stuhec@gmail.com

Most people like 3D visualizations. Whether it is in movies, holograms or games, 3D (literally) adds an extra dimension to conventional pictures. However, 3D data and their visualizations can also have scientific archaeological benefits: they are crucial in removing relief distortions from photographs, facilitate the interpretation of an object or just support the aspiration to document archaeology as exhaustively as possible. Since archaeology is essentially a spatial discipline, the recording of the spatial data component is in most cases of the utmost importance to perform scientific archaeological research. For complex sites and precious artefacts, this can be a difficult, time-consuming and very expensive operation.
In this contribution, it is shown how a straightforward and cost-effective hard- and software combination is used to accurately document and inventory some of the cultural heritage of the Cres/Lošinj archipelago in three or four dimensions. First, standard photographs are acquired from the site or object under study. Secondly, the resulting image collection is processed with some recent advances in computer technology and so-called Structure from Motion (SfM) algorithms, which are known for their ability to reconstruct a sparse point cloud of scenes that were imaged by a series of overlapping photographs. When complemented by multi-view stereo matching algorithms, detailed 3D models can be built from such photo collections in a fully automated way. Moreover, the software packages implementing these tools are available for free or at very low-cost. Using a mixture of archaeological case studies, it will be shown that those computer vision applications produce excellent results from archaeological imagery with little effort needed. Besides serving the purpose of a pleasing 3D visualization for virtual display or publications, the 3D output additionally allows to extract accurate metric information about the archaeology under study (from single artefacts to entire landscapes).

Keywords: *Structure from Motion, computer vision, 3D model, digital documentation*

1 Introduction

In the fields of cultural heritage, archaeology, art history and museology, accurate registration and documentation methods are needed. Besides the spectral documentation (i.e. colour), the exact recording of the geometrical properties of artefacts (such as vases), structures (such as buildings) or even complete landscapes is of the utmost importance. An accurate geometrical representation in three dimensions (3D)

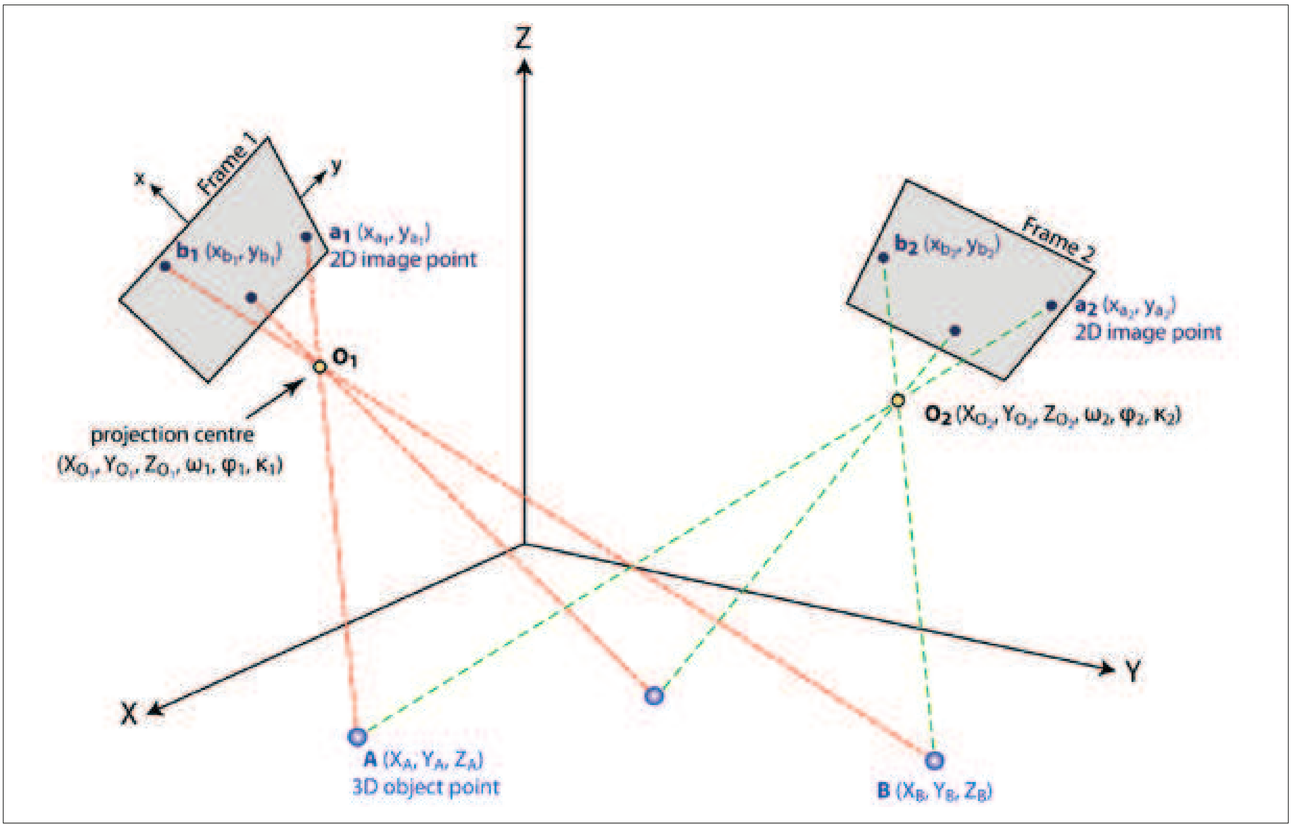

Fig. 1. Mapping of 3D object points onto 2D points in two frame images

allows to maximize the opportunities for future reproduction of the dataset, especially when it comes down to very fragile remains or objects that are subject to constant change (e.g. a ruin). Moreover, 3D models enable the extraction of information about dimensions and volume of these objects and structures, while also allowing for spatial analysis (in a landscape context) and digital sections or study the geometrical degradation over time.

However, accurate 3D documentation can be a difficult, time-consuming and very expensive operation. To date, different institutions and practitioners from a variety of scientific fields focus on the production and usage of digital 3D models. Many technologies and possible approaches exist (see Table 1), but all are characterised by certain disadvantages. For the documentation of structures and landscapes, standard survey methods using a differential GNSS (Global Navigation Satellite System) receiver or total station are often too time-consuming and generally do not provide the required data density to enable a highly detailed archaeological investigation. Airborne and terrestrial laser scanning or a photographic approach using systematic vertical aerial coverage certainly offer a viable approach to the problem, but the data acquisition is quite expensive while the processing often necessitates a thorough technical understanding and appropriate (i.e., costly) software. To document artefacts, often very accurate (and thus expensive) scanners are used, although many archaeologists still resort to more traditional techniques such as manual and less accurate drawings (Table 1). In this contribution, it is shown how a straightforward and cost-effective computer vision combination can be used to record and archive cultural heritage in four dimensions: 3D geometry and colour. The approach is very scalable, as it can be used for very small objects as well as complete landscapes. Moreover, the comparison with a laser scan of a small artefact will show how accurate it can be. Finally, the appealing visual character of the 3D models allows for a higher public participation and awareness of the archaeological heritage.

Method	Cost	Time	3D accuracy	Small artifacts	Large structures	Portable	Colour
Drawing	++	--	0	+	--	++	0
Total station	0	0	++	--	++	0	--
D-GNSS	-	+	+	--	++	0	--
Photographs	++	++	--	++	+	++	++
Laser scanning	--	++	++	++	++	-	+
This method	+	+	+/++	++	+	++	++

Table 1 - Comparison of different documentation methods

2 Method

Thanks to computer vision research in the last two decades, it is now possible to extract 3D geometry from a number of overlapping images. The approach is straightforward and requires no assumptions with regard to the camera parameters, extensive photogrammetric and computer vision knowledge of the user or the geometrical properties of the scenes. Moreover, simplicity is combined with geometrical quality due to the fact that the interior camera calibration parameters are automatically computed and a dense 3D model is extracted. In the next sections, the individual steps of this method will be detailed.

2.1 Structure from Motion (SfM)

Many tools and methods exist to obtain information about the geometry of 3D objects and scenes from 2D images. One of the possibilities is to use multiple image views from the same scene. Using techniques such as triangulation, an image point occurring in at least two views can be reconstructed in 3D (Figure 1). However, this can only be performed if the projection matrices of the images are known. These express the interior calibration parameters of the camera used (the focal length, the principal point location plus lens distortion coefficients) and camera poses (i.e., the location of the projection centre and the image orientation defined by the six exterior orientation parameters) at the moment of image acquisition. In computer vision, these orientation parameters are usually combined in the so-called projection matrices of the images[1], which can be determined by an approach called Structure from Motion (SfM)[2]. During this approach the relative projection geometry of the images is computed along with a set of 3D points that represent the scene's structure. SfM only requires corresponding image features occurring in a series of overlapping photographs captured by a camera moving around the scene.[3] Sometimes, this approach is also referred to as Structure and Motion, since both the structure of the scene and the motion of the camera (i.e. the different camera positions during image acquisition) are recovered.

In order to achieve this, SfM relies on algorithms that detect and describe local interest points (i.e. image locations that are in a certain way exceptional and are locally surrounded by distinctive texture) for each image and then match those 2D interest points throughout the multiple images (Figure 2A-C). This results in a number of potential correspondences (often called tie points). Using this set of correspondences as input, SfM computes the locations of those interest points in a local coordinate frame and produces a sparse 3D point cloud that represents the geometry/structure of the scene (Figure 2D). As mentioned previously, the camera pose and internal camera parameters are also retrieved.[4] There is thus no real need to use calibrated cameras and optics during the image acquisition stage, which makes the procedure very flexible and well-suited for almost any kind of imagery. One should only make sure that the images feature a sufficient overlap (around 60 % to 80 %) and it is preferable for every image to be captured from a unique location. Panning from the same location should thus be avoided.[5] Moreover, the objects being photographed

[1] Robertson, Cipolla 2002.
[2] Ullman 1979.
[3] Fisher et al. 2005; Quan 2010; Szeliski 2011.
[4] Hartley, Zisserman 2003; Szeliski 2011.
[5] Tingdahl et al. 2012.

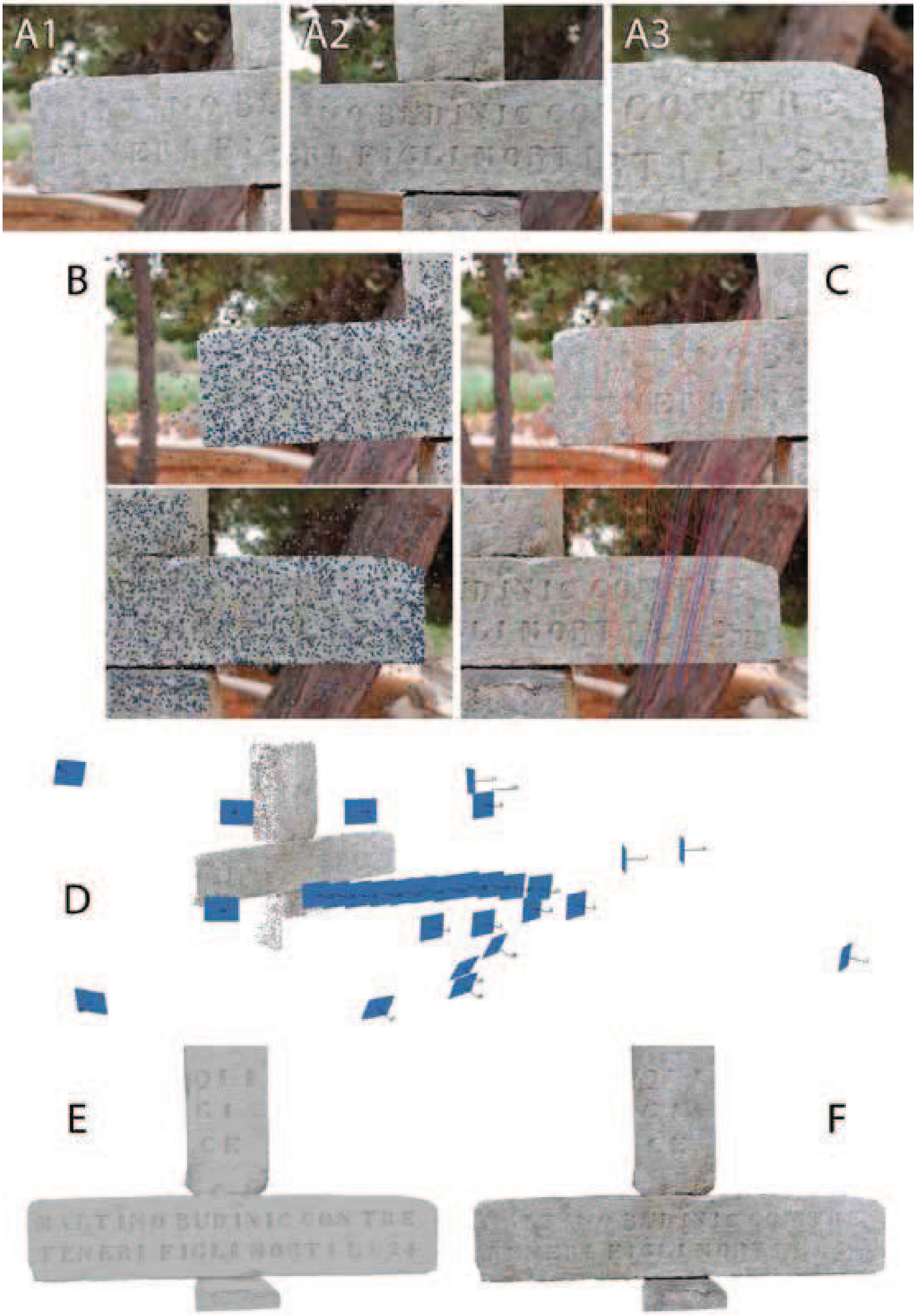

Fig. 2. (A1-3) three out of thirty photographs from a cross at Sv. Petar; (B) indicates the detected interest points from two of those images and (C) the matches found between those points. (D) shows the sparse point cloud and the camera positions provided by the SfM solution. Using MVS, a dense and accurate 3D model (point cloud or mesh) is created (E), which can afterwards be textured (F).

Company	Software	Free	SFM	MVS	Web	Orthophoto
Agisoft LLC	PhotoScan Standard		X	X		
Agisoft LLC	PhotoScan Professional		X	X		X
Matis Laboratory (I.G.N.)	Apero	X	X			
Matis Laboratory (I.G.N.)	MicMac	X		X		X
Univ. of Washington & Microsoft Corp.	Bundler	X	X			
Microsoft Corporation	PhotoSynth	X	X		X	
University of Washington	VisualSFM	X	X	X		
AutoDesk	123D Catch	X	X	X	X	
KU Leuven	Arc3D	X	X	X	X	
Eos Systems Inc.	PhotoModeler Scanner		X	X		X
Univ. of Illinois & Univ. of Washington	PMVS2	X		X		
3Dflow SRL	3DF Samantha	X	X			
Henri Astre & Microsoft Corp.	PhotoSynth Toolkit	X	X	X		
CTU Prague	CMPMVS	X	X	X	X	X
Acute3D	Smart3DCapture		X	X		

Table 2 - Some commercial and freely available SfM and MVS packages

need to possess sufficient unique texture. As a result, reconstructing a white wall without any texture will not work.

2.2 Multi-View Stereo (MVS)

After the SfM step, a sparse 3D reconstruction of the scene is generated. "Sparse" because it is only based on the reconstructed set of interest points. However, with the now known orientation of the images it becomes possible to create a texture-mapped dense 3D model and even compute orthophotographs. The essential step in this process is the computation of this denser 3D model. Alternatively, one could interpolate the sparse set of 3D points, but this would yield a far from optimal result. Therefore, it is better to run a Multi-View Stereo (MVS) algorithm to compute a dense estimate of the surface geometry of the observed scene (Figure 2E). Because these solutions operate on pixel values instead of on feature points[6], this additional step enables the generation of detailed 3D meshed models (or dense point clouds) from the initially calculated sparse point clouds, hence reproducing fine details present in the scene. Afterwards, this 3D model can be textured (Figure 2E) or used to generate orthophotographs.

As a result, this method largely accounts for most relevant kinds of geometrical degradations (geometric errors induced by the optics, the topographical relief and the tilt of the camera axis) and is capable of generating 3D models and orthophotos that are perfectly suited for archaeological purposes. Further, only minimal technical knowledge and user interaction are required. Finally, this approach can also work in the total absence of any information about the instrument the imagery was acquired with, although it is still advised to have at least information about the focal length of the camera applied. The extra investments needed for software and computing hardware are recovered easily when taking the time and cost savings of map production into account.

3 Software

In recent years, SfM has received a great deal of attention due to the free Bundler[7] and Microsoft's Photosynth[8]. To date, several SfM-based packages can be applied to obtain a (semi-)automated processing pipeline for image-based 3D visualisation. Often, these packages are complemented by an MVS approach (Table 2). The case studies presented in the next section have all been computed with PhotoScan

[6] Scharstein, Szeliski 2002; Seitz et al. 2006.

[7] Snavely et al. 2010.

[8] Microsoft Corporation 2010.

Professional. The choice for this software was based on its features, cost and completeness: it is currently the only commercial package that combines both SfM and MVS algorithms while additionally offering tools for generating orthophotographs, texture mapping and 3D post-processing.[9]

4 Case studies

SfM-based applications started to find their way into archaeological research about ten to fifteen years ago.[10] During the decade that followed, the SfM concept and dense matching techniques made great improvements and became capable of orienting very large datasets and delivering satisfactorily accurate dense 3D models.[11] The case studies described below show the potential of this combined SfM + MVS method using a variety of archaeological objects and structures.

4.1 Lady from Čikat

A small prehistoric figurine – the so-called "Lady from Čikat" – has proved to be a suitable object to test the accuracy of the proposed image-based method (Figure 3).[12] First, the figurine was scanned using the NextEngine laser scanner: a triangulation scanner that projects several laser stripes on the object and registers the position of the laser points by observing it with two cameras. Although it is a rather low-cost 3D scanner, the achievable geometric accuracy is generally around 0.5 mm.[13] To this end, the scanner is a very cost-effective means to digitise small archaeological objects.[14] After scanning in the highest accuracy macro mode, the different scans were aligned and cleaned in Geomagic Studio 2012 to yield the final laser scanned 3D model.

Besides 3D geometry, the scanner also captures 1.3 megapixel digital images to texture the scan. In this case, twenty of these low-quality images (Figure 3) were used as input to calculate a second 3D model using SfM and MVS. A thorough comparison between the two models revealed that the average deviation was only 0.16 mm (Figure 3). Given that this number is lower than the accuracy achievable with the NextEngine laser scanner, the image-based model can be considered equally accurate as the scanned model, even though it was calculated out of small and low-quality images.

4.2 Osor – city wall

To model a small part of the city wall around the ancient city of Osor,[15] first some general overview photographs were obtained. These are used to align the subsequently photographed details. For the final 3D model, only the detail photographs are used, as they allow to extract most accurate geometrical information. Although the view on the wall was obstructed by cars and containers (Figure 4), these were excluded from the final textured 3D model.

4.3 Osor – the Franciscan Monastery of Tertiary Glagolitic Monks and the Church of St. Mary of the Angels

Around the ruins of the Franciscan monastery and church,[16] about 1000 digital images were obtained (Figure 5-A). Although this image set constitutes a very nice photographic record of these buildings, they also showcase two shortcomings: first, no images were obtained above the buildings. As a result, the roofs cannot be decently reconstructed. To this end, some pole or UAV (unmanned Aerial Vehicle) imaging would be needed. Second, such a big image set necessitates a very powerful computer with a lot of RAM (Random-Access Memory) to process. Figure 5-B shows a section through the church. This model is not a full meshed model, but only a very dense point cloud. Anyhow, the amount of visible detail is striking. Moreover, this section showcases that it is possible to document both the inside and outside of a building in one model.

[9] Agisoft LLC 2012.
[10] E.g. El-Hakim et al. 2003; Pollefeys et al. 1998; Pollefeys, van Gool 2002; Pollefeys et al. 2003.
[11] Barazzetti 2011.
[12] Blečić Kavur 2012.
[13] Guidi et al. 2007; Polo, Felicísimo 2012.
[14] E.g. Hermon et al. 2012; Štuhec 2012; Tucci et al. 2011.
[15] Faber 1980.
[16] Fučić 1949.

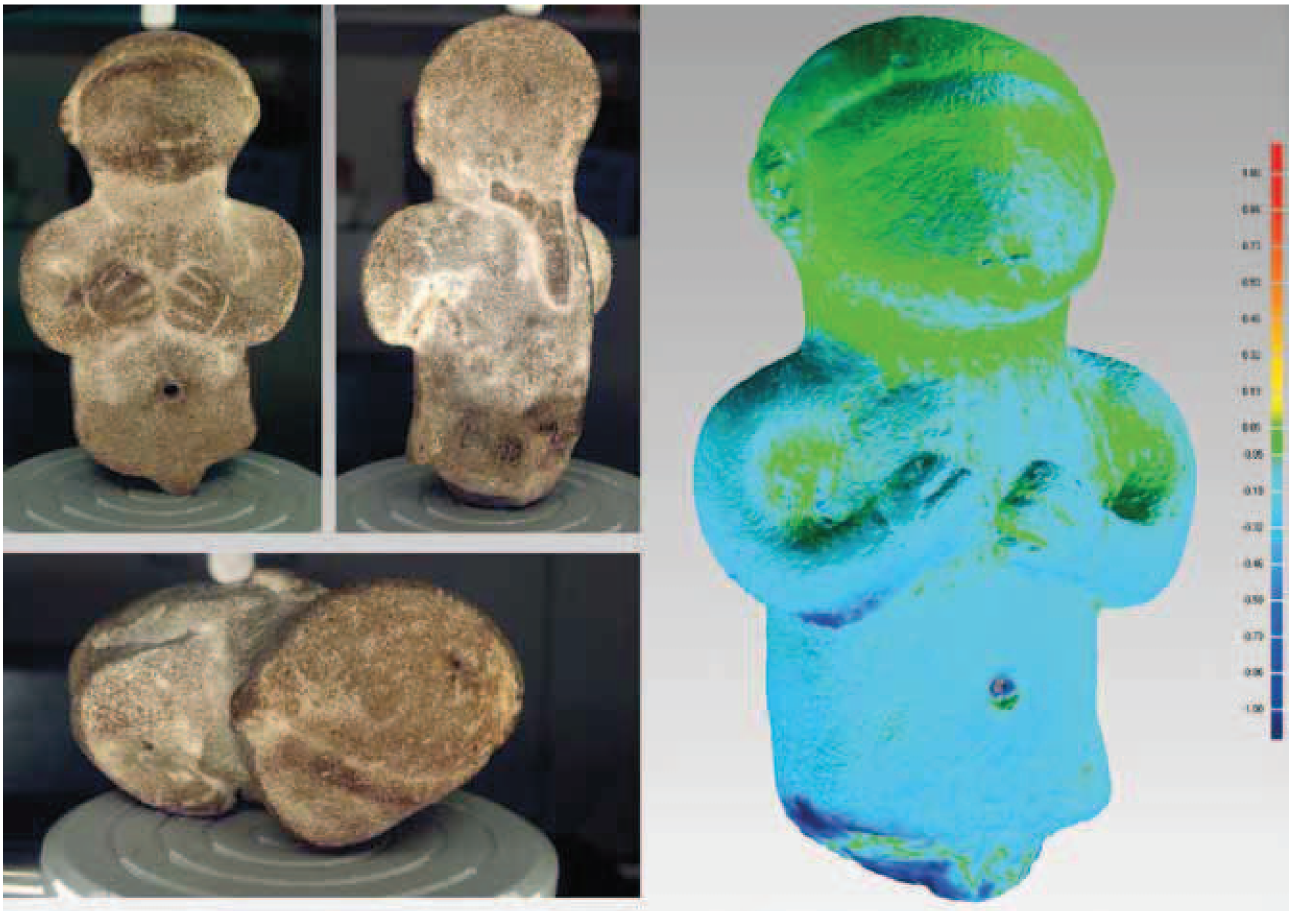

Fig. 3. Three of the twenty NextEngine images (on the left) from which a 3D model was calculated. On the right, the image-based 3D model is compared with the laser scanned model. The legend indicates the deviation between the two models in millimetres.

4.4 Benedictine Monastery of Saint Peter (Sv. Petar)

Spread over two days, the four walls of the old monastery at the small island of Sv. Petar[17] were photographed from the outside (Figure 6-A). As two sides of this building are in very close proximity to trees, many images (650 in total) had to be acquired to cover these walls with sufficient image overlap. Additionally, three tape measurements were collected (Figure 6-B). After scaling the 3D model to its real-world dimensions (Figure 6-C) with one of these measurements, the other two measurements were used as a simple validation. Comparing it with the distance measured on the 3D model resulted in a maximal difference of only 2 cm. Figure 6-D and E show a small part of the complete wall and its corresponding 3D model. Notice the amount of detail that can be extracted from the photographs.

[17] Bully, Čaušević-Bully 2012.

5 Conclusion

Straightforward 3D and even 4D documentation is very important in cultural heritage and archaeology. In this article, computer vision algorithms (Structure from Motion and Multi-View Stereo) were exploited to present an integrated, cost-effective, semi-automated approach to generate 4D models (geometry and colour) of archaeological frame images. This approach is straightforward and requires no assumptions with regard to the camera projection matrix or extensive photogrammetric and computer vision knowledge of the user. Only a digital camera and a decent computer are needed, while the user has to make sure that enough overlapping and sharp images are acquired. Moreover, simplicity is combined with geometrical quality due to the fact that the inner camera calibration parameters are automatically computed. Furthermore, the approach is suited for both small objects

Figure 4 - The city wall of Osor as coloured 3D model, showing that obstacles such as cars and containers were filtered out.

and large structures or even complete landscapes.[18] Finally, this approach can also work in the total absence of any information about the instrument the imagery was acquired with, although it is still advised to have at least information about the focal length of the imaging system applied.

Of course, it is not all roses. First, the processing is very computer resource intensive while the method is not applicable for the individual image. At least two – but preferably more images – are needed for accurate 3D computation. In addition, erroneous alignment of the imagery can occur when dealing with very large photo collections, images that suffer from excessive noise or blur, highly oblique photographs or photographs that have a very dissimilar appearance. Finally, the accuracy of the final products and the recovered camera parameters is often less than results yielded by the expensive and rigorous photogrammetric approaches. However, differences are often small while the approach presented here is superior in versatility and flexibility. The latter point cannot be over-estimated, as many archived images do not fulfil the constraints (e.g. camera parameters) that are essential for the more standard photogrammetry approaches. As a result, the presented approach can also be used to extract accurate 3D models from old frame imagery.

[18] In the case that one has aerial imagery such as in Verhoeven et al. 2012.

Copyright

Acknowledgements

The Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology (archpro.lbg.ac.at) is based on an international cooperation of the Ludwig Boltzmann Gesellschaft (A), the University of Vienna (A), the Vienna University of Technology (A), the Austrian Central Institute for Meteorology and Geodynamic (A), the office of the provincial government of Lower Austria (A), Airborne Technologies GmbH (A), RGZM-Roman- Germanic Central Museum Mainz (D), RAÄ-Swedish National Heritage Board (S), IBM VISTA-University of Birmingham (GB) and NIKU-Norwegian Institute for Cultural Heritage Research (N). This article has been written within the framework of the Austrian Science Fund (FWF): P 24116-N23. Finally, we would like to thank Zrinka Ettinger Starčić for her support and access to the sites and artefact.

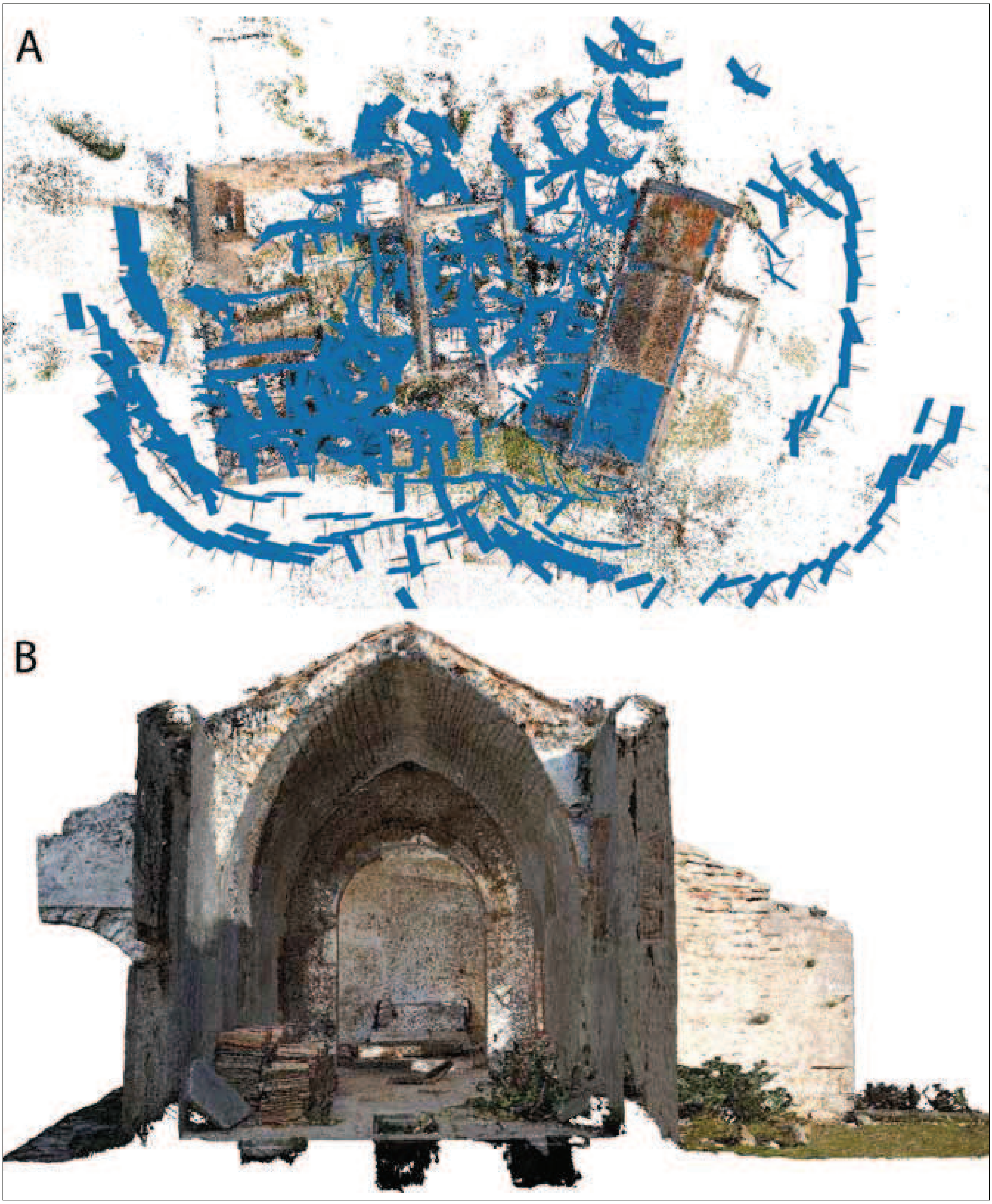

Figure 5 – (A) The locations of the 1000 images from which the Monastery and church were documented; (B) a cut through the dense point cloud representing the church.

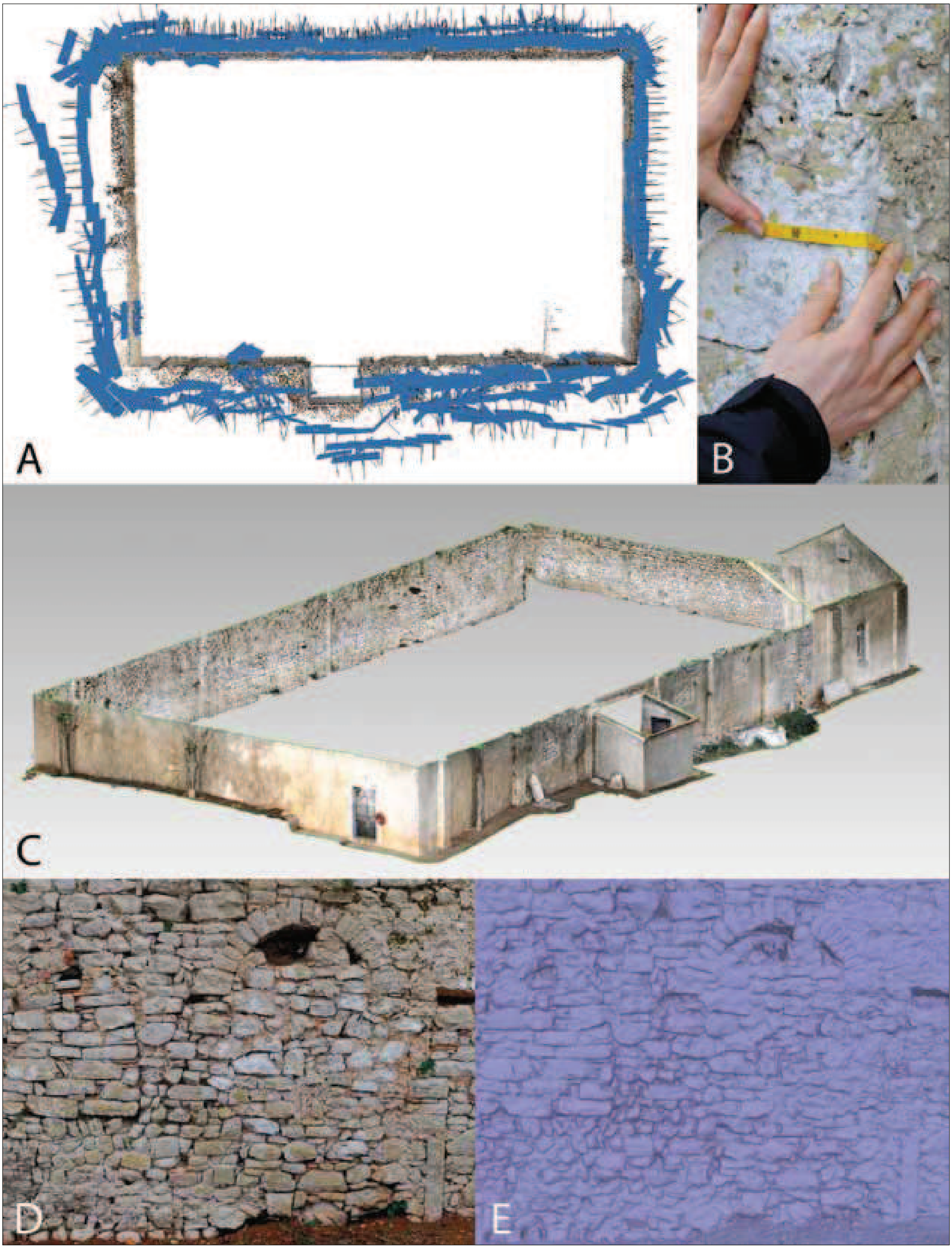

Figure 6 – (A) All 650 images acquired to document the Early Christian Basilica; (B) measurements were taken for scaling and validation; (C) the final scaled model; (D) a detail of a photograph and its corresponding 3D model.

Geert Verhoeven, Nives Doneus, Michael Doneus, Seta Štuhec

OD PIKSELA DO MREŽE – PRECIZNO I JEDNOSTAVNO 3D DOKUMENTIRANJE KULTURNE BAŠTINE CRESKO-LOŠINJSKOG ARHIPELAGA

(Sažetak)

3D vizualizacija je vrlo omiljena. Je li riječ o filmovima, igrama ili hologramima, 3D (doslovno) daje dodatnu dimenziju ustaljenim slikama. Međutim, 3D podaci i njihova vizualizacija također imaju prednosti u arheologiji: oni su ključni za bolje fotografije reljefa, olakšavaju interpretaciju arheoloških objekata ili jednostavno podržavaju težnju arheologa za što boljom dokumentacijom kulturne baštine. Kako je arheologija u suštini disciplina koja se bavi „prostorom“, evidentiranje takvih podataka je od velikog značaja za znanstvena istraživanja. Kod kompleksnih nalazišta i vrijednih objekata, dokumentacija se vrlo lako može pretvoriti u težak, dugotrajan i vrlo skup proces. U ovom prilogu je prikazano jednostavno i ekonomično rješenje za preciznu dokumentaciju kulturne baštine na primjeru nekih objekata i građevina na cresko-lošinjskom arhipelagu. Koristeći različite metode, pokazan je put do izvrsne 3D vizualizacije, koja omogućava preciznu dokumentaciju, bez obzira radi li se o pojedinačnim objektima ili o arheološkim krajolicima.

Ključne riječi: trodimenzionalna rekonstrukcija prostora iz slika (*Structure from Motion*), računalni vid, 3D model, digitalna dokumentacija

REFERENCES

Agisoft LLC 2012
Agisoft PhotoScan User Manual: Professional Edition, Version 0.9.0. Agisoft LLC, 49 pp. http://downloads.agisoft.ru/pdf/photoscan-pro_0_9_0_en.pdf. Date of access: 13 February 2013.

Barazzetti 2011
L. Barazzetti, *Automatic Tie Point Extraction from Markerless Image Blocks in Close-Range Photogrammetry*, http://www.dirap.unipa.it/autec/uploads/NeHtaDUab5DEBgFTaRstFSPtc8P9GUFkYdDZHqZJ.pdf

Blečić Kavur 2012
M. Blečić Kavur, *Dama iz Čikata (The Lady from Čikat)*, Arheološki vestnik 63, 2012, 57–77.

Bully, Čaušević-Bully 2012
S. Bully, M. Čaušević-Bully, *Saint-Pierre d'Ilovik. Une station maritime majeure du nord de l´Adriatique, de l´Antiquité au Moyen-Âge*, Histria antiqua 21, 2012, 413–426.

El-Hakim et al. 2003
S.F. El-Hakim, J.-A. Beraldin, M. Picard, *Effective 3D Modeling Of Heritage Sites*, in: *4th International Conference 3-D Digital Imaging and Modeling, Banff, Canada. October 6-10*, 2013, 302–309. http://www.3dphotomodeling.org/El-Hakim-3DIM_03.pdf

Fisher et al. 2005
R.B. Fisher, K. Dawson-Howe, A. Fitzgibbon, C. Robertson, E. Trucco, *Dictionary of computer vision and image processing*, Wiley, Chichester 2005.

Fučić 1949
B. Fučić, *Izvještaj o putu po otocima Cresu i Lošinju*, Ljetopis Jugoslavenske Akademije 55, 1949, 31–76.

Guidi et al. 2007
G. Guidi, F. Remondino, G. Morlando, A. Del Mastio, F. Uccheddu, A. Pelagotti, *Performances Evaluation of a Low Cost Active Sensor for Cultural Heritage Documentation*, in: *Proceedings of the VIII Conference on Optical 3D Measurement Techniques. Volume 2, Zürich, Switzerland. 9-12 July 2007, ISPRS*, 59–69. http://3dom.fbk.eu/sites/3dom.fbk.eu/files/pdf/guidi_etal_O3D07.pdf

Hartley, Zisserman 2003
R. Hartley, A. Zisserman, *Multiple view geometry in computer vision, 2nd ed*, Cambridge University Press, Cambridge2003.

Hermon et al. 2012
S. Hermon, D. Pilides, G. Iannone, R. Georgiou, N. Amico, P. Ronzino, *Ancient Vase 3D Reconstruction and 3D Visualization*, in: *Revive the Past. Proceeding of the 39th Conference on Computer Applications and Quantitative Methods in Archaeology, Beijing, Japan. 12-16 April, 2011*, Pallas publications, Amsterdam 2012, 59–64.

Microsoft Corporation 2010
Photosynth. Microsoft Corporation. http://photosynth.net/

Pollefeys et al. 1998
M. Pollefeys, R. Koch, M. Vergauwen, L. van Gool, *Virtualizing Archaeological Sites*, in: *Proceedings of the 4th International Conference on Virtual Systems and Multimedia (VSMM) 98*, 1998. http://www.cs.unc.edu/~marc/pubs/PollefeysVSMM98b.pdf

Pollefeys, van Gool 2002
M. Pollefeys, L. van Gool, *Visual modelling: from images to images*. The Journal of Visualization and Computer Animation 13 (4), 2002, 199–209, DOI: 10.1002/vis.289.

Pollefeys et al. 2003
M. Pollefeys, L. van Gool, M. Vergauwen, K. Cornelis, F. Verbiest, J. Tops, *3D recording for archaeological fieldwork. IEEE Computer Graphics and Applications 23 (3)*, 2003. 20–27, DOI: 10.1109/MCG.2003.1198259. http://www.cs.unc.edu/~marc/pubs/PollefeysCGA03.pdf

Polo, Felicísimo 2012
M.-E. Polo, Á.M. Felicísimo, *Analysis of Uncertainty and Repeatability of a Low-Cost 3D Laser Scanner*, Sensors 12 (7), 2012, 9046–9054. DOI: 10.3390/s120709046.

Quan 2010
L. Quan, *Image-based modeling*, Springer, New York 2010.

Robertson, Cipolla, 2002
D.P. Robertson, R. Cipolla, *Building architectural models from many views using map constraints, in: Computer Vision - ECCV 2002, 7th European Conference on Computer Vision, Copenhagen, Denmark. May 2002*, Springer, Berlin-Heidelberg, 2002, 155–169.

Snavely et al. 2010
K.N. Snavely, I. Simon, M. Goesele, R. Szeliski, S.M. Seitz, *Scene Reconstruction and Visualization From Community Photo Collections. Proceedings of the IEEE 98 (8)*, 2010, 1370–1390. http://www.cs.cornell.edu/~snavely/publications/papers/proc_ieee_scene_reconstruction.pdf

Štuhec 2012
S. Štuhec, *Dvoinpolrazsežnostna (2,5D) in trirazsežnostna (3D) vizualizacija artefaktov*, Master, Ljubljana 2012.

Szeliski 2011
R. Szeliski, *Computer vision: Algorithms and applications. Texts in Computer Science*, Springer, New York 2011.

Tingdahl et al. 2012
D. Tingdahl, M. Vergauwen, L. van Gool, *ARC3D: A Public Web Service That Turns Photos into 3D Models*, in: F. Stanco, S. Battiato, G. Gallo (Eds.), *Digital Imaging for Cultural Heritage Preservation. Analysis, Restoration, and Reconstruction of Ancient Artworks. Digital Imaging and Computer Vision Series*, CRC Press, Boca Raton, 2012, 101–125.

Tucci et al. 2011
G. Tucci, D. Cini, A. Nobile, *Effective 3D digitization of archaeological artifacts for interactive virtual museum*, in: *3D-ARCH 2011: "3D Virtual Reconstruction and Visualization of Complex Architectures". Proceedings of the 4th ISPRS International Workshop, Trento, Italy. 2-4 March 2011*, ISPRS, 2011, 413–420. http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXVIII-5-W16/413/2011/isprsarchives-XXXVIII-5-W16-413-2011.pdf

Ullman 1979
S. Ullman, *The Interpretation of Structure from Motion*, Proceedings of the Royal Society B. Biological Sciences 203 (1153), 1979, 405–426. DOI: 10.1098/rspb.1979.0006. http://dspace.mit.edu/bitstream/handle/1721.1/6298/AIM-476.pdf?sequence=2

Verhoeven et. al. 2012.
G. Verhoeven, M. Doneus, C. Briese, F. Vermeulen, *Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs*, Journal of Archaeological Science 39 (7), 2012, 2060–2070. DOI: 10.1016/j.jas.2012.02.022.