PH06
INVESTIGACIÓN

3D modelling in cultural heritage using structure from motion techniques

José Manuel Pereira Uzal **01|**

The evolution of photogrammetric techniques thanks to the development in the computer vision techniques and strategies as the structure from motion (SFM). This strategies are a new generation of highly versatile and accessible tools for the professional in documentation and heritage preservation, thanks to low implementation costs. Throughout this article, we will study the evolution of 3D technology as well the fundamentals of SFM and the potential use of this strategies in documentation tasks in cultural heritage how a way for the recording of global or particular damage.

Keywords
Photogrammetry | Digital image | Cultural heritage | Heritage preservation | Structure from motion (SFM) | 3D |

Modelado 3D en patrimonio cultural por técnicas de *structure from motion*

José Manuel Pereira Uzal **01|**

El actual desarrollo de las técnicas de fotogrametría, gracias a la evolución en el campo de la visión por computador y en particular las estrategias de structure from motion (SfM), se nos plantea como una nueva generación de herramientas altamente polivalentes y accesibles al profesional de la documentación o protección del patrimonio, gracias a los bajos costes en implementación. A lo largo de este artículo se explora la evolución de las tecnologías 3D, así como los fundamentos del SfM y el potencial uso de éste en las tareas de documentación de bienes culturales en general, o alteraciones en particular.

Palabras claves
Fotogrametría | Imagen digital | Patrimonio histórico | Preservación patrimonio cultural | Structure from motion (SFM) | 3D |

URL <http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

Salvo que se indique lo contrario, esta obra está bajo una licencia Reconocimiento -No Comercial- Sin obra Derivada 3.0 España Creative Commons
Documento traducido gracias a la desinteresada colaboración de Morote Traducciones. Artículo publicado inicialmente en español: ph investigación n.º 1 (diciembre, 2013)

01| DigitalHeritage editor (http://www.jpereira.net/)

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp.49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

INTRODUCTION

Human beings' eagerness to portray the three dimensions is almost as old as our artistic manifestations themselves and, although a subject of study from the Classical Period until the Renaissance, it would be in 1840 when Charles Wheatstone invented the stereoscope; this was capable of recreating –with an apparent simplicity– the sensation of depth in pairs of pictures or figures in which slight differences between the two resulted in the perception of depth when processed by our visual system. Around 1858, years after the invention of the daguerreotype, this photographic process began to present itself as an ideal means for the recreation of stereoscopic scenes which would later gain relevance as a tool for the study of different disciplines. It would be precisely around the same time when the German architect Albrecht Meydenbauer (ALBERTZ, 2002) began to shape the concept and techniques of photogrammetry as a tool for the study of the geometric properties of objects and scenes from photographs.

Meydenbauer's achievements should not only be regarded in terms of the development of photogrammetry, but we should also consider him one of the pioneers of the graphic documentation of heritage for its preservation, as his objectives included the use of photographic images with the aim of preserving building and monument geometry to then be able to reconstruct them in the event of a catastrophe.

Over the years to come the principles of photogrammetry would be applied to the development of all types of devices such as the stereo-autograph and the photocartograph, inspired precisely by the anaglyph process patented by Louis Ducos du Hauron in 1891. Anaglyph or anaglyphic stereoscopic images are those that are formed by the famous pairs of images coloured in red and blue. By using glasses with red and blue gel filters –and thanks to the phenomenon of human vision of additive colour synthesis– the aforementioned dominant colours are partially neutralised and a three-dimensional image is perceived in black and white.

Nowadays, after decades of evolution, photogrammetry, whether for distant objects –such as aerial photogrammetry– or for nearby objects –intended for the study of scenes or objects close to the observer– has shifted from complex mechanical devices to comfortable computer tools. Some of these tools, such as the well-known PhotoModeler from the company Eos Systems, have been offering solutions in the field of photogrammetry –for over 20 years– as a strategy for 3D modelling and the geometric study of nearby scenes and objects.

Nonetheless, advances in artificial vision in recent decades have introduced new approaches to 3D modelling to the market such as

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp. 49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

structure from motion (SfM), based on the phenomenon by which human and animal visual systems can reconstruct three-dimensional structures from 2D images projected onto the retina thanks to the movement of these structures as perceived by the observer, or the movement of the observer with regard to these structures. This phenomenon can also be described by the term kinetic depth, perhaps more widely used in the fields of visual perception, while SfM frequently appears in association with the field of computational vision. Along these lines, commercial products such as Photoscan from Agisoft or early tools as 123Catch, or modern 3D modelling tools as REMAKE (formerly known as Memento) from the Autodesk giant or modern products as Capturing Reality, allow for the modelling and study of 3D geometry in a very intuitive and simple way.

Although there were already studies on SfM in the 1980s, this last decade has led to consistent tools that are capable of solving complex 3D models fairly quickly and effectively thanks, in part, to the increase in the computing power of multi-core processors. We are referring to contributions such as that of Changchang Wu with Multicore Bundle Adjustment (CHANGCHANG; AGARWAL; CURLESS et al. 2011) and creator of the free tool VisualSFM. This software also draws on the work of Yasutaka Furukawa –Google Maps engineer–, developer of CMVS/PMVS tools that are responsible for generating dense point clouds from large image collections; or the work of Noah Snavely, which gave rise to the Bundler free projects, based on the classical Levenberg-Marquardt algorithm (LOURAKIS; ARGYROS, 2005); and the work of Yasutaka. All of these projects have created an important substrate with the contributions of a good number of algorithms and free 3D modelling tools using 2D images that are open and available to any user.

While photogrammetry required a previous characterisation of the cameras to be used for capturing the images, and in many cases these shots had to be somewhat planned, SfM is more spontaneous and does not need any sort of planning or calibration of the cameras. It is even possible to use images taken with different cameras at different points in time; however, it is still necessary to locate overlapping areas or areas in our objects or scenes that are included in different images.

Therefore, it is complex and confusing to speak of SfM and photogrammetry separately nowadays (MUNDY, 1993), as the majority of photogrammetry tools frequently end up incorporating SfM processes in order to automate certain routines and free the user from the tedious tasks of manually locating the common points among different images, thus being able to utilise large amounts of images to describe, in as much detail as possible, the geometry of scenes and objects.

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp.49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

The possibilities of new tools based on SfM, or on modern photogrammetry techniques, provide us with an excellent approach to 3D graphic torno a 1858, años después de la invención del daguerrotipo, dicho proceso fotográfico comienza a presentarse como un medio ideal para la recreación de escenas estereoscópicas, las cuales ganarán relevancia como herramienta para el estudio de diferentes documentation (STANCO, 2011; WULFF, 2010). Given their accessibility –from a financial point of view– and ease of implementation or ease of use, these new tools are an alternative to 3D modelling methods based on laser technology (laser imaging detection and ranging, LIDAR) or structured light. They are not only inaccessible to many professionals owing to the price of acquiring them, but they are also complicated to use and adapt to the workflows characteristic of the documentation and conservation of heritage properties; on many occasions they are poorly understood from a topography point of view.

METHODOLOGY

The result of a photogrammetry or SfM process is initially a discrete point cloud which can vary in density depending on the connections detected among the images. In the traditional processes of manual adjustment of common points among images, these clouds were not especially populous; with SfM, however, these clouds can end up having thousands of points. These discrete, or not very dense, clouds are the result of what is known as the Bundler adjustment.

This concept is closely tied to photogrammetry and allows for the spatial repositioning of a series of overlapping points between images as well as the positions of the cameras which took said images with regard to the scene. These points can be manually established or, in the case of SfM, automatically detected using what is known as the SIFT process, or scale-invariant feature transform (LOWE, 1999; MAKADIA, 2007), by which common points or characteristics are detected between pairs of images, thus allowing us to compare hundreds of images in order to extract large amounts of common or key points.

Although these discrete clouds can yield information about geometry, they are inadequate for the thorough assessment of an object or scene; thus, the final step in these SfM processes tends to be a dense point cloud with millions of points that describe the surfaces and geometry of objects in greater detail.

When we talk about dense point clouds, whether they are obtained using SfM or photogrammetry, we are referring to a set of vertices described in an XYZ three-dimensional coordinate system. In addition to the spatial

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp. 49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

information, in the case of SfM and some LIDAR, each vertex or point is accompanied by a colorimetric description in the RGB model.

This combination of geometric or spatial information with colorimetric data is especially interesting when it comes time to compiling descriptive information about a scene or a piece of work. During the SfM processes this colorimetric information is extracted from the pixels of the images used in the process; thus, if these images are colorimetrically reliable this will be transferred to our dense point cloud. This phenomenon resulting from the 3D modelling therefore requires that we pay special attention to the management of colour in the images used in order to ensure colorimetric accuracy to the extent possible.

Once we have obtained our dense point cloud we have a document on which to carry out both geometric and colorimetric estimates; additionally, we have a place where we will be able to connect our workflow with virtualisation tasks for informational purposes only, where photo-realistic textures and synthetic lighting take precedence over the accuracy of the models.

All 3D models –obtained via one technique or another– can be assigned a scale based on a known distance between two points which allows us to carry out measurements in a particular system of units. However, we can also contextualise 3D models in a coordinate system by assigning control points (ground control points, GCP). This situation allows us to create scenarios with different models and connect them to geographical information systems, create layer-based models and, in particular, make estimates between different point clouds using tools such as CloudCompare; this tool allows us to compute the distances between points on two seemingly similar clouds with the aim of revealing and quantifying, through false colour maps, the possible global or particular differences among models. With these types of methods we can document and quantify structural alteration processes caused by either dimensional changes or material losses.

To evaluate the potential of these tools in heritage, in particular as a means for the documentation or evolution of alterations, a natural geological setting was searched for with a rock substrate of limestone with possible material losses due to flaking. These natural alterations will be exploited to document a situation of material loss.

CASE STUDY

Our workflow begins by taking 6 photographs with a common 10 Mp DSLR camera; said photographs will be used for the SfM process.

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp.49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

Next, within the VisualSfM tool, we will sequentially carry out the steps described in the methodology section relating to the construction of a 3D model based on SfM techniques:

1. Detection of characteristics is carried out using the SIFT algorithm. At this point a description of the relevant characteristics of each image is created, regardless of their scale or the technical aspects of the image itself such as brightness, colour, contrast, etc. This description is stored in a database in order to be used in the following point.

2. From the descriptions of the characteristics of each image, pairs of images are compared and the candidates for the best overlapping or close points are located for each pair of shots.

3. After locating the common characteristics between pairs of images, the Bundler adjustment is carried out; this will produce the dispersed point cloud which will be more or less dense depending on the number of common points among the images from the previous step.

4. Finally, the dense point cloud is generated from the SIFT data and the Bundler adjustment, creating our finished model which is described by a dense point cloud of some 552,963 points for an approximate area of 50x50 cm.

After conducting this process with the two models both before and after causing a loss of material (images 1-4), we can now use the models to calculate the distance between their points with the CloudCompare tool.

For the purpose of computing said distances between two models, the models must be “registered” using an algorithm known as iterative closest point (ICP) so that both models go on to occupy the same space from which we can compute the differences. This aspect is critical since we could generate false differences owing to defects in the alignment of both models (images 5 and 6).

Thanks to the calculation of the distances between clouds and the false colour maps, we can reveal non-overlapping aspects between pairs of models with relative ease. Although this method goes beyond simple documentation and virtualisation, it is presented as one of the largest potentials for the 3D modelling of objects, artwork and structures.

DISCUSSION

During the project “4D Rock art, Monitoring and Preventive Conservation of the Rock Art of the Mediterranean Basin on the

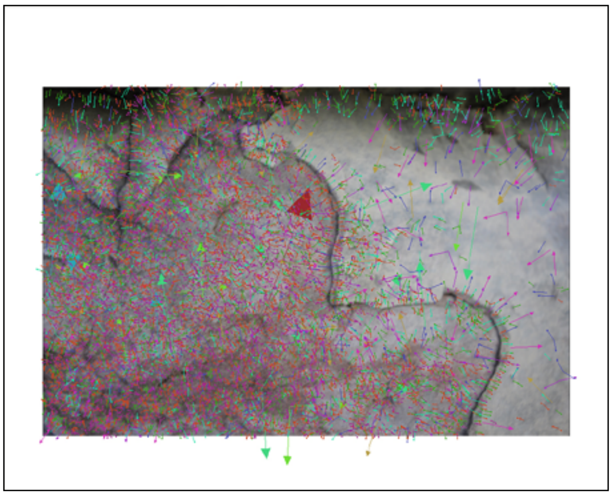

Image 1 |
Description of the relevant characteristics of each image

Image 2 |
Once the characteristics of each image are known, they can be located between pairs

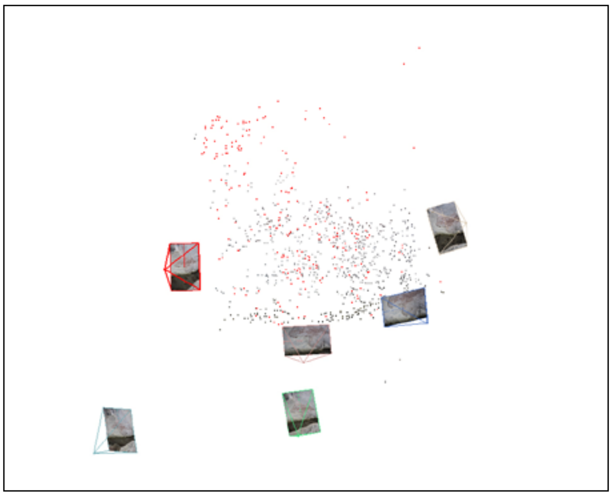

Image 3 |
Appearance of the dispersed point cloud following a Bundler adjustment

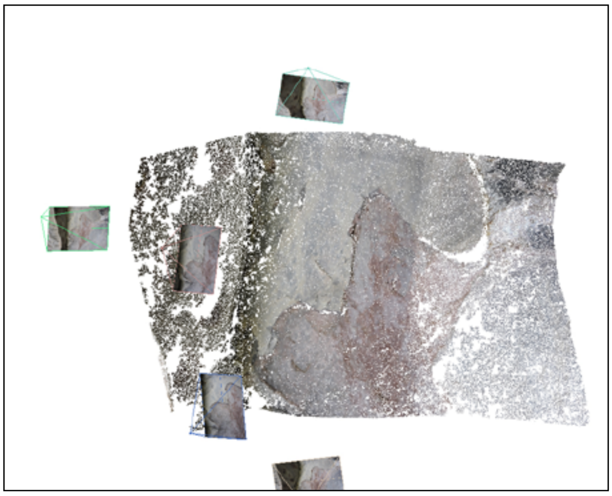

Image 4 |
Dense point cloud with the locations of the cameras or points of view of the shots

Iberian Peninsula" (Spanish Ministry of Education, Culture and Sport, 13 March 2013), these types of methodologies were intensively implemented and models between 3 and 14 million points were obtained with the aim of revealing potential alterations with the passing of time. These models require between 60-100 photographs to obtain complete models of caves with an important level of detail, thus entailing a greater demand for calculation and time resources in order to successfully conclude each project.

With regard to the field of digital preservation, although many manufacturers of software and tools strive to contribute their own file formats and viewers, the reality is that we currently have a certain

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp.49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

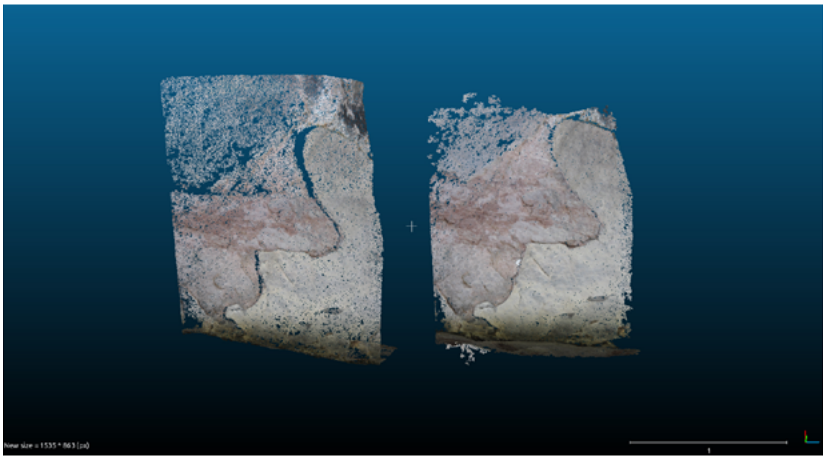

Image 5 |
Both models uploaded to CloudCompare before superimposing them in order to compare them

Image 6 |
Result of the calculation of distances between both models, where the area missing in the second model was determined from the distance between points

number of file formats encoded in ASCII or XML that can facilitate our access to said documentation at a future point in time.

This way, file formats such as the stanford triangle format (.ply) or the wavefront file (.obj) allow us to describe our models based on point clouds using readily understandable ASCII files, as the information is described in rows and columns using the format X Y Z R G B. Along similar lines, we have file formats such as X3D or COLLADA which are based on open standards encoded in XML. Although the file descriptions based on plain text, ASCII and XML are not ideal compared to binary formats, they are a very good alternative in terms of the preservation or conservation of digital information in the long run.

Nowadays it apparently is not easy to reproduce 3D content on our computers because of the unusual nature of these file formats; however, the reality is far different since, for some time now, we have had some open source and multiplatform viewers-editors for 3D

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp. 49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

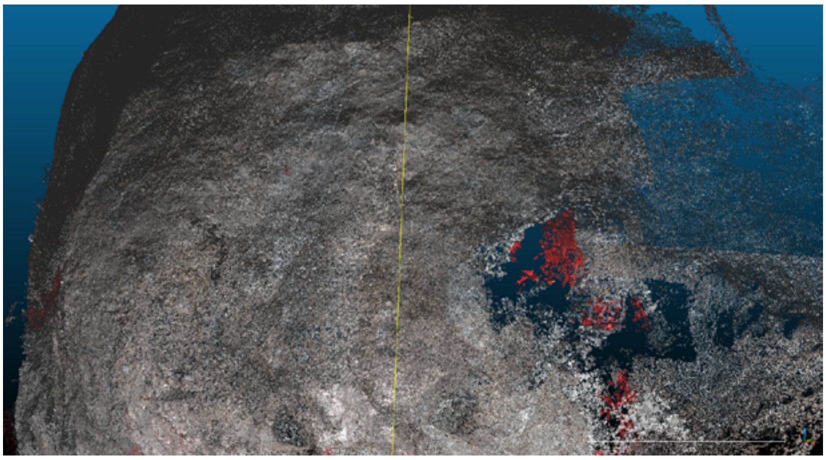

Image 7 |
3D modelling of the Solana de Cobachas cave 6 (Nerpio, Albacete)

Image 8 |
Detection of an increase in vegetation owing to a difference between two models

models of nature, such as the aforementioned CloudCompare, the popular MeshLab, and the support offered by Adobe Photoshop to 3D models encoded in wavefront file, one of the most popular three-dimensional data exchange formats described in ASCII, or universal 3D file, a recent standardisation resulting from the partnership of several manufacturers and encoded in binary.

There is also the Mac Os X native support from the pre-view tool (from the Snow Leopard version onwards) for file formats based on COLLADA –another recent, open-standard format that is described in XML.

CONCLUSIONS

SfM techniques are proposed as a major step forward in the documentation of heritage, not only for their important levels of geometric precision (VERHOEVEN, 2012; ARIAS, 2006; WULFF,

PEREIRA UZAL, J. M. (2016)
3D Modelling in Cultural Heritage using Structure from Motion Techniques
ph investigación [en línea], n.º 6, junio de 2016, pp.49-59
<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/135>

2010), but also for the accessibility (DONEUS, 2011; REUA, 2012) and directness of this technique, whose levels of detail are solely conditioned by the resolution and number of images to be used.

SfM as an open-source tool allows us to not only bring 3D modelling closer to almost any professional in the documentation or protection of heritage, but it also lets us tackle 3D modelling jobs in places that are not easily accessible for instruments based on LIDAR or structured light. It also offers us immediacy (VERHOEVEN, 2012) and spontaneity throughout documentation tasks since it requires very little planning when it comes time to taking photographs, compared to the tedious planning of capturing photos in traditional photogrammetry, or the inoperability of certain computers in the face of certain working conditions.

Additionally, although the results of SfM are dense point clouds, it is very easy to connect it to other workflows based on the virtualisation of objects or scenes via the regeneration of netted surfaces and photo-realistic textures, as well as the creation of orthophotos or orthoimages free of perspective errors.

BIBLIOGRAPHY

ALBERTZ, J. (2002)
Albrecht Meydenbauer-Pioneer of Photogrammetric Documentation of the Cultural Heritage. En 18th International Symposium of CIPA. *The photogrammetric record.; International Committee for Architectural Photogrammetry; Potsdam, Germany, 2001; Sep, 2002, 505.* Potsdam, Germany: Remote Sensing and Photogrammetry Society, 2002, p. 19 y ss.

ARIAS, P.; ARMESTO, J.; DI-CAPUA, D. et al. (2006)
Digital photogrammetry, GPR and computational analysis of structural damages in a mediaeval bridge. *Engineering Failure Analysis*, vol. 14, n.º 8, December 2007, pp. 1444-1457

BRUTTO, L. M.; MELI, P. (2012)
Computer Vision Tools for 3D Modelling in Archaeology. *International Journal of Heritage in the Digital Era*, vol. 1, suplemento 1, pp. 1-6

CHANGCHANG, W.; AGARWAL, S.; CURLESS, B. et al. (2011)
Multicore bundle adjustment. En *Procs. IEEE Conf on computer vision and pattern recognition*, 2011, pp. 3057-3064

DONEUS, M.; VERHOEVEN, G.; FERA, M. et al. (2011)
From deposit to point cloud: a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. *Geoinformatics*, n.º 6, 2011, pp. 81-88

HARTLEY, R. I.; MUNDY, J. I. (1993)
Relationship between photogrammetry and computer vision. *Proc. SPIE* 1944, Integrating Photogrammetric Techniques With Scene Analysis and Machine Vision, 92, September 24, 1993

LOWE, D. (1999)
Object recognition from local scale-invariant features. En *The proceedings of the seventh IEEE international conference on computer vision.* sl: IEEE Computer Society, 1999, 2 vol., pp. 1150-1157

LOURAKIS, M. I. A.; ARGYROS, A. A. (2005)
Is Levenberg-Marquardt the Most Efficient Optimization Algorithm for Implementing Bundle Adjustment? En *Proceedings of the Tenth IEEE International Conference on Computer Vision*, 2005, vol. 2, pp. 1526-1531

MAKADIA, A.; GEYER, C.; DANIILIDIS, K. (2007)
Correspondenceless Structure from Motion. *International Journal of Computer Vision*, 75 (3), 2007, pp. 311-327

REU, J. DE; PLETS, G.; VERHOEVEN, G. et al. (2013)
Towards a three-dimensional cost- effective registration of the archaeological heritage. *Journal of Archaeological Science,* vol. 40, n.º 2, February 2013, pp. 1108-1121

STANCO, F.; BATTIATO, S.; GALLO, G. (ed.) (2011)
Digital imaging for cultural heritage preservation: analysis, restoration, and reconstruction of ancient artworks. Boca Raton, Fla.; London: CRC, 2011 (Digital Imaging and Computer Vision)

VERHOEVEN, G. (2011)
Taking Computer Vision Aloft-Archaeological Three-dimensional Reconstructions from Aerial Photographs with PhotoScan. *Archaeological Prospection*, n.º 18(1), pp. 67-73

VERHOEVEN, G.; DONEUSB, M.; BRIESEC, CH. et al. (2012)
Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science,* n.º 39, 2012

WULFF, R. (2010)
Image-based 3D Documentation in Archaeology. En *DAGM 2010 32nd Annual Symposium of the German Association for Pattern Recognition* [en línea] <http://resources.mpi-inf.mpg.de/conferences/dagm/2010/pdfs/wulffdagm2010.pdf> [consulta: 23/06/2016]

Interesting web links

- 123Catch: http://www.123dapp.com
- Agisoft: http://www.agisoft.ru
- Bundler: http://phototour.cs.washington.edu/bundler/
- CloudCompare: http://www.danielgm.net/cc/
- CMVS/PMVS: http://www.di.ens.fr/pmvs/
- MeshLab: http://www.meshlab.net
- Photomodeler: http://www.photomodeler.com
- Proyecto 4D arte rupestre, monitorización y...: http://www.4darterupestre.com/
- VisualSFM: http://homes.cs.washington.edu/~ccwu/vsfm/
- Autodesk Remake: https://memento.autodesk.com
- Capturing Reality: https://www.capturingreality.com