3D documentation in archaeological fieldwork: a case study from the site of Metsamor

Otto Bagi
Independent researcher

Abstract: Three-dimensional recording techniques, although growing rapidly in efficiency and applicability for archaeologists, have still not been turned to full account, mainly because they require for the most part expensive equipment and know-how. In this respect, photogrammetry is unique, being relatively cheap and easy to use. The joint Armenian–Polish archaeological project in Metsamor drew on the possibilities of this technology, which is relatively new to archaeological proceedings, in the daily recording system used at the site. The following article discusses the step-by-step application of photogrammetry in field documentation and its positive impact on archaeological work, Metsamor being taken as a case in point.

Keywords: Metsamor, three-dimensional registration, 3D, photogrammetry, digital documentation method, digital archaeology

The Early Bronze Age–Iron Age III archaeological site of Metsamor in Armenia, excavated since 2013 by a joint Polish–Armenian team, was targeted from the start as a testing ground for various new methods and technologies to improve the efficiency of archaeological fieldwork, while keeping costs at bay. The site, which lies on the outskirts of modern-day Taronik, some 35 km west of Yerevan, is a large citadel on a hilltop and a lower town on terraces situated around the hill (for recent archaeological fieldwork, see Jakubiak 2017, in this volume).

A previous case study (De Reu, De Clercq et al. 2013) demonstrated that implementing three-dimensional recording methods in the daily system of documenting archaeological fieldwork can considerably speed up the process of drawing plans, enhancing at the same time their accuracy. Moreover, excavation is in essence a destructive process, hence the importance of recording as much information as possible during fieldwork. Three-dimensional models of an excavated area capture not just the surface geometry in previously unimaginable detail, but also the surface colors of the site, making it easier to detect soil discolorations and patterns. Repeating the process at regular intervals will create a series of models that will encapsulate the progress of the excavation. Taken together, the models provide archaeologists with an extensive and mobile — meaning easy to transport and available at any time for reinterpretation — backup of the work carried out through

Fig. 1. Sequence of selected three-dimensional models from Metsamor sector VIII square S16 showing the progress of excavation in the 2015 season (3D models and visualization O. Bagi)

the season(s), something previously impossible [*Fig. 1*]. At Metsamor, the key issue was to overhaul the conventional paper-and-pencil documentation system and replace it with digital plans and profile drawings based on orthophotos[1] generated from three-dimensional surface models.

METHODS

In general, there are two ways of obtaining an accurate image of a surface without perspective distortion, and both involve generating three-dimensional surface models. The use of laser scanners is a somewhat more sophisticated method, but for most excavation projects the high cost and the need for specially trained staff to operate the equipment are restrictive. The Metsamor team needed a substitute technology that was more readily available and, most importantly, more user-friendly. Photogrammetry, in conjunction with the low-cost Agisoft Photoscan software package, offered a potent alternative for a fraction of the cost of laser scanners, considering that earlier studies (Doneus et al. 2011; Grussenmeyer et al. 2008) have demonstrated its accuracy to be comparable to laser-scanning.

The main advantage of photogrammetry is that it requires neither special equipment nor extensive expertise in order to achieve highly accurate results. Therefore, its application costs are significantly lower than in the case of other three-dimensional recording methods. A digital camera of any kind suffices (Callieri et al. 2011: 36), although a full-frame DSLR camera equipped with a wide-angle lens[2] is preferred. With the application of advanced computer vision algorithms embedded into Photoscan, such as structure-from-motion (SfM), it is possible to obtain three-dimensional data from a sequence of two-dimensional pictures taken in the field by a digital camera. The concept of the algorithm is based on the automatic detection of feature points and tracking of their movement throughout a sequence of

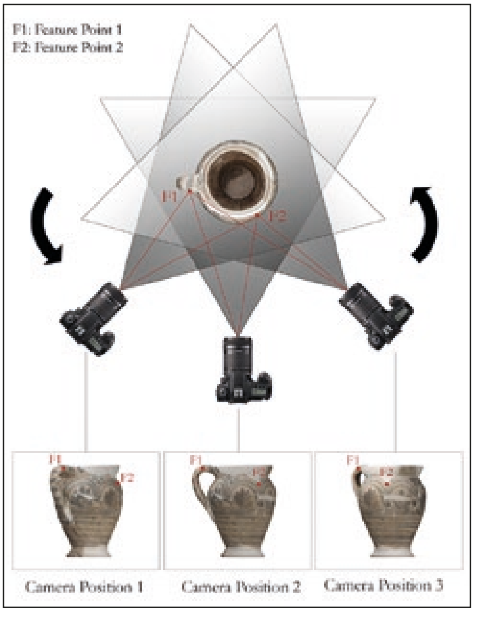

Fig. 2. Concept of feature point tracking and three-dimensional registration used by the structure-from-motion algorithm (Photo O. Bagi)

[1] "Orthophoto offers a geometrically correct image in which all possible deformations, due to camera tilt and variations in object height, are corrected" (De Reu, Plets et al. 2013: 1111).

[2] 18 mm or lower focal length lenses offer a substantial field-of-view even in combination with crop-sensor cameras, although ultra wide-angle objectives, such as fish-eye lenses should be avoided because of their extreme image distortion.

overlapping pictures taken from different camera positions [*Fig. 2*]. With the help of collected data it is possible to determine not only the exact location of each camera, but also the precise location of each feature point in three-dimensional space (Westoby et al. 2012: 301–302). The collected information, in combination with data retrieved by multi-view stereo reconstruction algorithms, allows the software to recreate the geometry of an object with high precision (Verhoeven 2011: 68). Moreover, the addition of manually established ground control points (GCP),[3] measured by a total station in the field, allows archaeologists to check the accuracy of the three-dimensional model. Besides, knowing the exact location of these marker points allows the software to automatically scale the model to its real-world dimensions and align it to the established coordinate system used for the excavation project's documentation system. Therefore, it is possible to perform precise measurements directly on the three-dimensional model without the need of visiting the excavation field.

Since the workflow within the software is highly automated, the level of human intervention needed to attain the final result is minimal. Thus, it provides an extremely user-friendly experience, one that does not require extensive knowledge in order to achieve sufficient results. However, experience and practice, especially in the field of photography, can significantly shorten the processing time, and increase the accuracy and quality of the final result.

WORKFLOW

In principle, the workflow of the whole process can be broken down into three main stages: photographing, photogrammetric processing and drawing. The first is just two steps, the others comprise several sub-steps. Importantly, in opposition to the conventional paper-and-pencil documentation system, photographing is the only phase that takes place in the field. This allows archaeologists to complete most of the documentation work indoors, thus minimizing mistakes due to fatigue and increasing the possibility of multi-tasking due to the highly-automated nature of the photogrammetric process.

PHOTOGRAPHING

Before the photographing starts, GCPs have to be placed and measured in the documented area. A minimum of three of these points should be present in the photographed area (Verhoeven et al. 2013: 47), but it is advisable to aim for more (at least 5–6) in order to have sufficient backup GCPs in case of error. Their placement is completely arbitrary, but it is good practice to disperse them evenly across the area of interest.

The photogrammetric process, as said above, is based on a sequence (or several sequences) of photos taken from various camera angles and positions covering the whole area of interest. The way of acquiring these images can have significant impact on the time spent in the field as well as on the speed of photogrammetric processing and the quality and accuracy of the final result. Therefore, this stage is considered the most

[3] A number of highly visible marker points placed within the photographed area and measured along the three axes (x,y,z) by a total station or GNSS RTK.

important within the whole documentation process (Kjellman 2012: 20).

In the case of a surface, whether excavation trench or architecture, there are several possible approaches to obtaining the photos needed for the three-dimensional documentation. The photographer can either use ground pictures taken by a hand-held or mounted camera (on a tripod or a photography mast), an UAV,[4] or a combination of the two. The highest resolution and quality are obtained, not surprisingly, by the first method thanks to the proximity of the camera to the subject. However, it also produces the highest number of images because of the smaller field-of-view of a camera held closer to the ground. More photos are needed to cover the whole area of interest and this can significantly increase the time spent in the field as well as the size of data generated and the speed of photogrammetric processing. Mounting a camera high on a tripod or mast will lower orthophoto quality, but will increase the camera's field-of-view substantially, accelerating therefore the process of photographing by decreasing the number of pictures required to cover the subject of the documentation. Undeniably, the deployment of UAVs provides the largest field-of-view and the fastest process of photographing, however this comes with a significant trade-off in image resolution and, in some cases, precision [*Fig. 3*].

As always, a middle way seems to be the best solution. Combining the methods, that is, using an UAV to photograph large swathes of horizontal surfaces, and a hand-held or mounted camera for important details and vertical or near-vertical surfaces,

Fig. 3. Two magnified fragments of orthophotos covering the same area: left, created from close-range images taken early in the morning and right, generated from photos taken in the afternoon at higher altitude with an UAV. Note the difference in sharpness and visibility (Photos M. Truszkowski, O. Bagi)

4 UAV stands for Unmanned Aerial Vehicle; this includes remotely operated multirotors, popularly known as drones, that are capable of carrying digital cameras.

can provide a flexible compromise between speed and quality that can be easily adjusted to fit individual scenarios [*Fig. 4*].

Naturally, working outdoors gives one little or no control over light conditions, yet the amount of shadow in the pictures can affect the quality of the orthophoto [see *Fig. 3*]. This in turn could have a negative effect on the legibility and interpretability of the acquired data. Therefore, photographing should ideally take place in cloudy weather, a rare instance on sites in the Middle East to start with. A suitable alternative is to photograph early in the morning when the whole trench is in the shade or at midday when shadows are the shortest.

In all cases, after manually establishing and measuring the ground control points, one has to aim at obtaining a sequence of overlapping and sharp[5] images parallel to the photographed surface(s) [see *Fig. 4*]. This can take an experienced photographer as little as 15–20 minutes in the case of a standard 5 m by 5 m trench depending on its depth and content. As discussed before, a higher number of pictures results in higher accuracy and quality, but also in longer processing time. Therefore, it is important to find a reasonable compromise between the two extremities. Beginnings can be difficult, but with practice and experience this will become considerably easier.

PHOTOGRAMMETRIC PROCESSING

Depending on the light conditions during a photographing session, an optional step can be included into the workflow before moving on to the photogrammetric processing. In the case of the presence of unwanted shadows, the images can be enhanced by increasing the brightness of the dark areas[6] using open source image manipulation software, such as GIMP. This can increase the quality of the final orthophoto substantially, helping at the same time the photogrammetry software to read as many pixels as possible in the sequence of images, which in turn can lead to increased accuracy.

Photo alignment

As mentioned before, Agisoft Photoscan workflow is divided into clearly distinguishable steps. The first step, after feeding the photos into the software, consists of determining the precise locations of the camera in the three-dimensional space relative to the photographed subject at the moment of taking each individual picture. This step is called photo alignment. Agisoft Photoscan utilizes the previously discussed fully automated structure-from-motion algorithm to register and track a set number of feature points. Therefore, human input needed during this initial step is minimal. Calculating camera locations is based entirely on the information collected automatically from the images by the software.

Upon finishing the photo alignment, the processed data is presented as a draft, a so-called spare point cloud [*Fig. 5* top]. This is essentially a loosely spaced group of points floating in the three-dimensional space that represents the approximate shape and color of the photographed subject.

[5] In order to ensure that all of the photographed area stays in focus, it is advisable to use a narrow aperture setting (f/18 or higher), while keeping ISO as low as possible (100–400) to avoid noise in the pictures.

[6] Compared to compressed JPEG, shooting pictures in the RAW file format makes the process of enhancing areas covered by deep shadows much faster and easier. To save storage space, images can be converted to JPEG following this step.

This allows archaeologists to have a general overview of the collected information and detect any errors that might have occurred either during the photographing or alignment process. The software offers an option to disable or realign any photos that are not correctly displayed. Moreover, using simple tools one can crop the area of

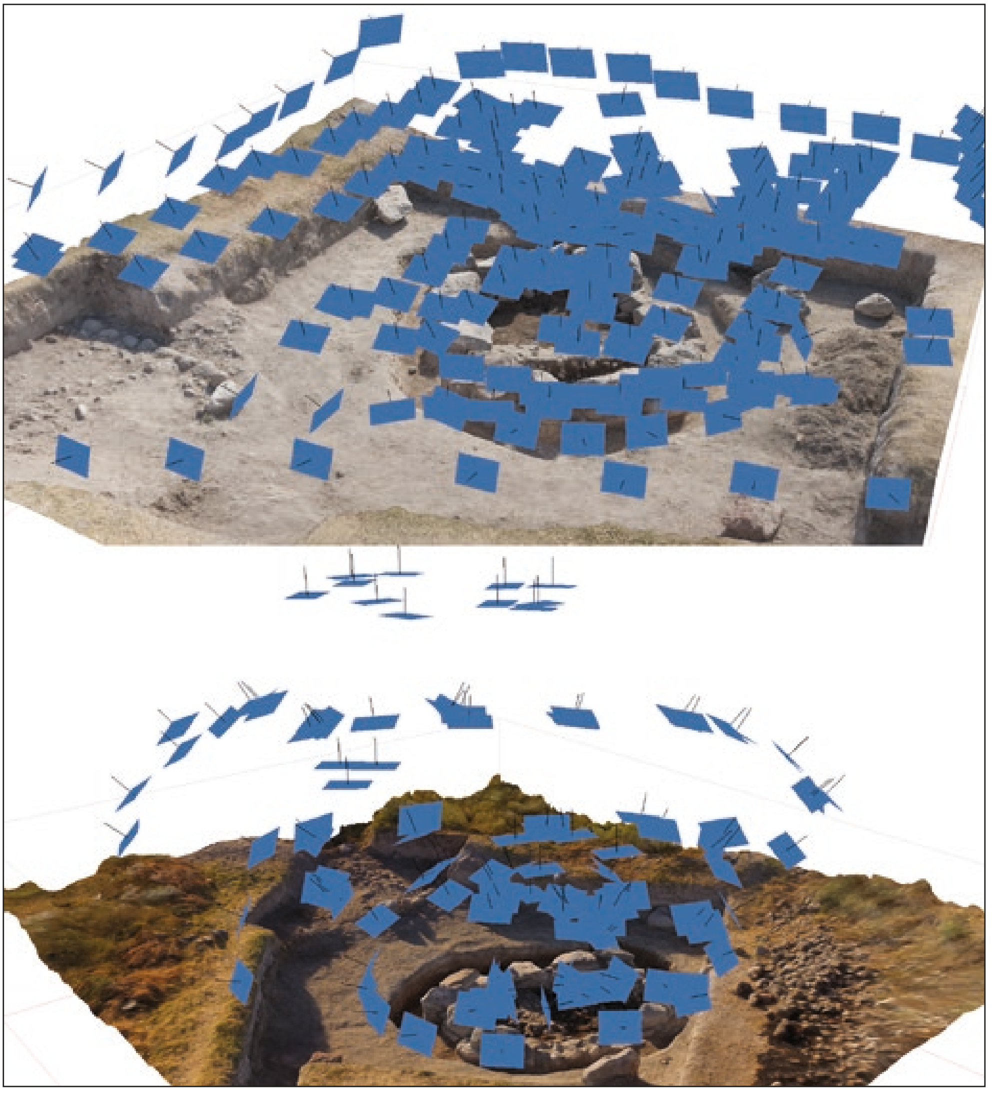

Fig. 4. Comparison of camera positions while capturing the same area with two different methods: top, only close-range images, and bottom, a combination of UAV and close-range photos. Note the difference in the number of pictures needed to cover the same area, 222 and 124, respectively (3D models and visualization M. Truszkowski, O. Bagi)

interest and delete all the unwanted noise generated during photo alignment before moving on to the next, more hardware-demanding step.

Dense point cloud generation

Dense point cloud generation is principally the resolution enhancement of the spare point cloud. The number of floating points is increased by the software from the range of tens of thousands to millions [see ***Fig. 5*** bottom]. Although the process is based on different algorithms, such as multiview stereo (Szeliski 2010: 558–570), it is just as fully automated as the one used during the previous phase. However, it is important to note that the process is considerably more hardware-deman-ding.

This leads to one genuine drawback of the photogrammetric documentation

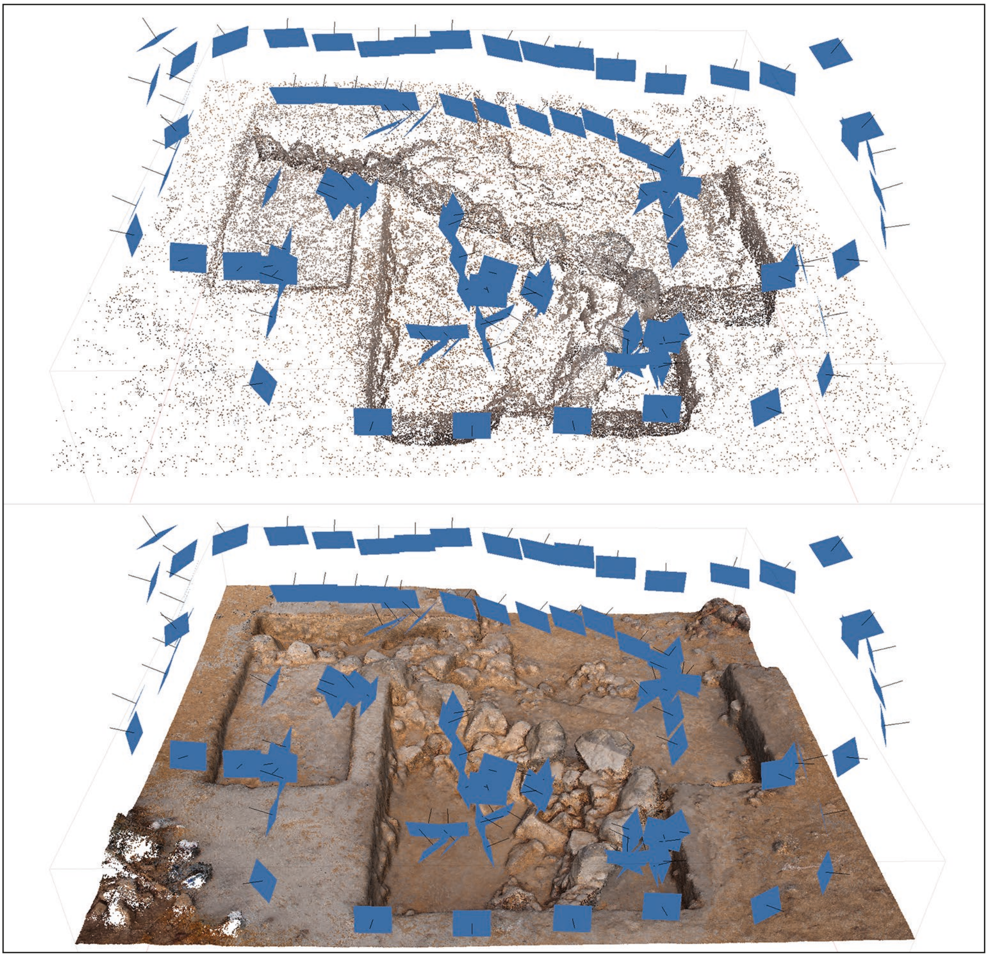

Fig. 5. Spare point cloud (top) and dense point cloud (bottom) of the same excavation trench containing 125.572 and 6.690.955 points, respectively (3D model and visualization O. Bagi)

method, namely the sheer amount of computing power required by the software for calculations. In order to reconstruct a subject in detail, Photoscan analyzes and attempts to find a match for all pixels in each consecutive image. Depending on the number and resolution of the photos this can be a cumbersome task requiring a powerful laptop[7] to complete the task within an acceptable (satisfactory) timeframe. As said, the approach used in the field during photographing can have a significant impact on the processing time, especially during this step. Therefore, it cannot be stressed enough that for efficient documentation work, it is essential to find the ideal method that will produce the lowest number of photos, while yielding a sufficiently detailed orthophoto adequate for plan drawing.

Mesh and texture generation

As the name suggests, a dense point cloud is a group of unconnected points organized in the shape of the photographed subject. In order to obtain a detailed three-dimensional copy of the target of the documentation, Agisoft Photoscan needs a solid surface, a so-called mesh [***Fig. 6***], upon which the texture[8] can be projected.

After optimizing the number of points, the mesh is, essentially, generated by connecting the remaining individual points (vertices) to form a pattern of triangles that describes the surface of the subject [***Fig. 7***]. The result can be displayed in three different ways: by showing only the connecting lines (edges); by displaying the uniformly colored triangular planes (faces) filling the space between points; or by rendering to each of the aforementioned triangular planes a rough estimate of their original color. These display modes are referred to as wireframe, solid and shaded [see ***Fig. 6***].

Upon completing mesh generation, a highly detailed geometry of the documented object is ready. However, at this point the color information is still rather sketchy. Texture is required to attain the real surface colors of the target in ultra-high resolution. Just as all the previous steps, this one is also highly automated in Agisoft Photoscan. After entering the initial resolution parameters,[9] the software automatically calculates the texture from the original photo sequence.

The mesh and texture calculations are somewhat less hardware-demanding than the first two steps and can be completed in a relatively short period of time.[10]

Reference points and orthophoto generation

Before being able to extract an orthophoto from a finished three-dimensional model, it needs to be scaled and aligned to its real-world size and orientation. In order to do so, the software needs external reference points with precise three-axial (x, y, z) measurements. Each previously established marker (prior to photographing) needs to

7 At least a higher mid-range or lower high-end laptop equipped with a dedicated video card and at least 16 GB of memory is recommended.

8 This is essentially a montage of photos created from the pictures taken in the field that is stretched upon the whole area of the three-dimensional model.

9 The recommended resolution for plan drawing is 8192 or 12,288 pixels width depending on the size of the documented area.

10 Based on personal experience, the processing time of mesh and texture generation is approximately ten times shorter than the combined calculation time of photo alignment and dense cloud generation.

be manually located and marked on the model [*Fig. 8*]. Once the task is done, the coordinates, measured earlier with a total station, can be entered into the project file. The inserted data is used by the software not only to automatically scale and position the model, but also to assess its accuracy. Agisoft Photoscan can provide accuracy measurements separately along the x, y and z axes (easting, northing and elevation), as well as the average of the three axes. If the result is not satisfactory, the program offers an option for optimization. This automated algorithm recalculates the camera locations taking into account the recently entered total station measure-

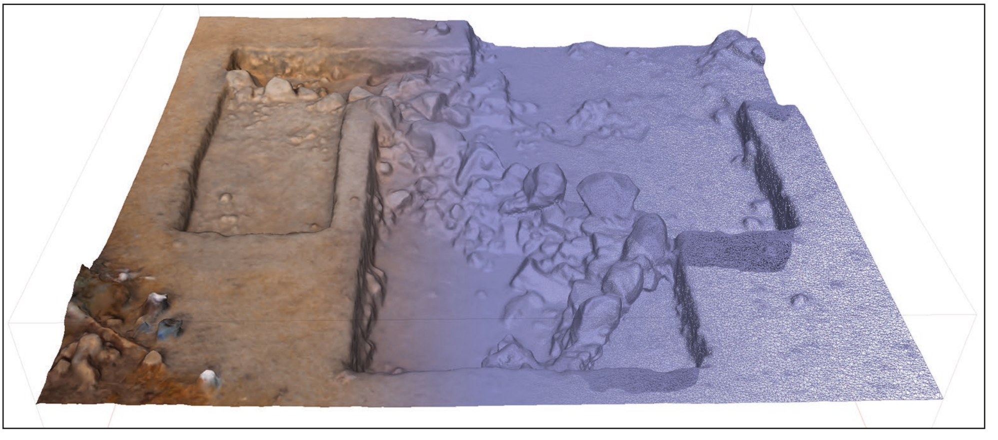

Fig. 6. Combined three-dimensional model of an excavation trench showing (left to right) the mesh in shaded, solid and wireframe modes (3D model and visualization O. Bagi)

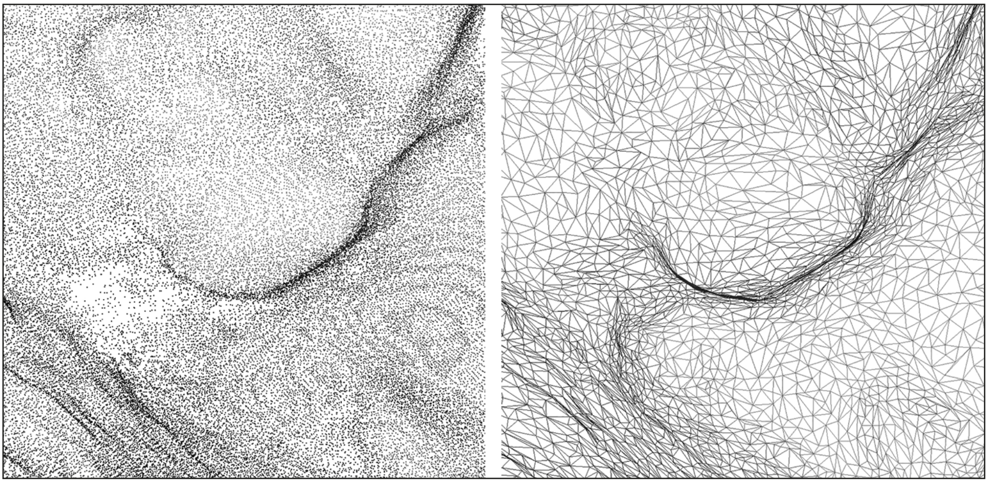

Fig. 7. Dense point cloud (left) and wireframe (right) of the same segment of a three-dimensional model (3D model and visualization O. Bagi)

ments. This process can significantly improve the accuracy of the geometry, however, it requires repeating of all the previous steps, which can consume valuable time.

Yet again, it cannot be emphasized enough that the way the photos were obtained in the field can have a significant impact on the accuracy of the 3D model. With practice and experience one can achieve precision measured in millimeters even in the case of larger excavation trenches.

From this point onwards the orthophoto generation process is fairly straigh forward. After entering the resolution parameters, the software calculates the horizontal (plan) or vertical (profile) orthographic projection of the documented object automatically. Moreover, the user may also generate, just as effortlessly, a digital elevation model (DEM) of the target with contour lines at pre-set intervals [*Fig. 9*]. This allows archaeologists to conduct basic measurements, such as elevation, distance, volume or even cross-section, with ease directly on the three-dimensional model. The finished orthophoto can be exported in three file formats; JPEG, TIFF or PNG.

Fig. 8. Reference point (plastic bottle cap) marked on the finished three-dimensional model (Photo O. Bagi)

DRAWING

Archaeologists have several software solutions at their disposal to render three-dimensional data into two-dimensional plans. Beside the most popular vector graphic programs, such as CorelDraw or Adobe Illustrator, there are some fairly competitive open source solutions, like Inkscape, that can provide a viable alternative for expeditions operating on a tight budget. Regardless of the software in use, the workflow during plan drawing is relatively similar in all cases. It can be broken down into two simple steps: scaling and drawing.

Scaling

The file formats offered for exportation by Agisoft Photoscan are widely recognized by most vector graphic programs. Therefore, it should not pose a problem to import the finished orthophoto into the drawing software used in the documentation system of the archaeological mission. The purpose of doing so is to use the image as a backdrop to the plan [*Fig. 10*]. This allows the user to trace the outline of each feature of interest, be it architecture, soil discoloration or any other element of the orthophoto worth recording. However, before one proceeds with drawing, the imported image has to be resized to match the scale used in the documentation.

It is highly advisable to use a uniform scale in the two-dimensional documentation throughout the season, which can make life much easier when one attempts to combine drawings from different areas. The only information required to resize the orthophoto to the desired scale is the known distance between at least two distinguishable points. There are various features in the image that can be used to this end, for instance the corners of a trench or the markers used during photographing. The required measurements can be easily obtained from Agisoft Photoscan.

Drawing

After successfully scaling the image, drawing is simply a matter of retracing the features visible in the orthophoto [see ***Fig. 10***]. This step is very much like digitizing hand-drawn plans, therefore should be familiar even to those who are accustomed only to pencil-and-paper documentation and have never used a fully digital recording system before.

Even though the final result might look similar to drawings done by traditional methods (after digitization), the crucial difference lies in the details. That is to say, the level of accuracy measured in millimeters and the amount of effort put into the workflow in order to obtain a plan

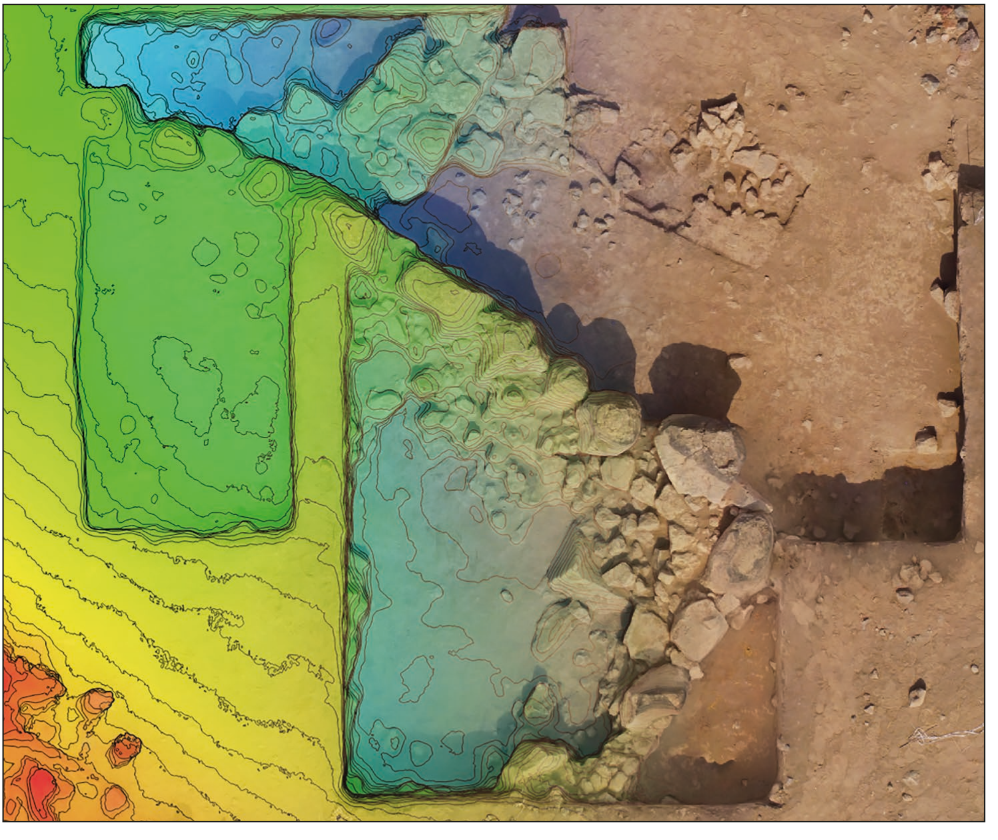

Fig. 9. Combined image of a final orthophoto (right) and DEM (left) with contour lines set to every 5 cm (Photo O. Bagi)

or section drawing. As a matter of fact, the whole process can be finished within a few hours.[11] This allows archaeologists to use digital recording on a daily basis, therefore making the excavation documentation not only more comprehensive, but also more useful in decision-making during the fieldwork.

CONCLUSIONS

Documentation techniques used in the field have changed little over the years and most excavation projects in the early 21st century are still relying on recording methods developed over a hundred years ago. With rapid development of photogrammetric software, three-dimensional documentation can offer a highly potent alternative to the traditional pencil-and-paper approach. It provides unmatched

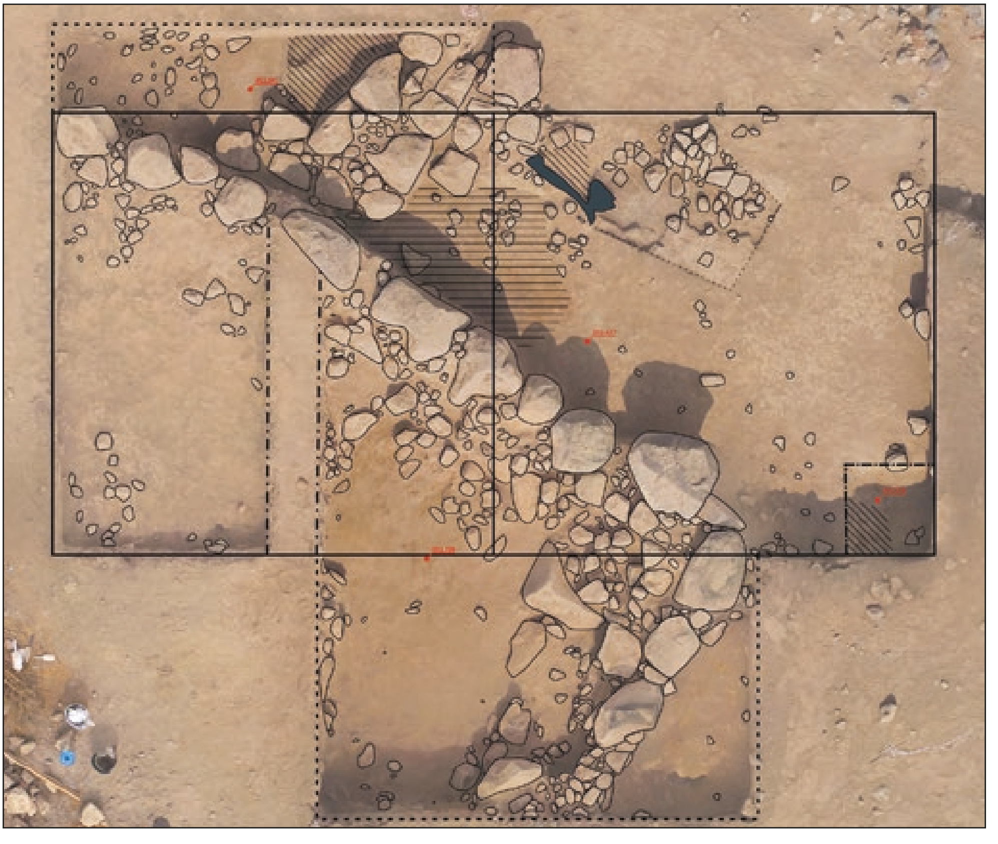

Fig. 10. Plan drawing based on an orthophoto placed in the background (Photo and drawing O. Bagi)

[11] Based on personal experience, for an advanced user the whole process of recording takes on average, depending on the scenario, approximately 2 to 4 hours for a 5 m x 5 m trench.

accuracy that can be achieved within a fraction of the time needed for hand-drawn plans.

The unparalleled speed of the process allowed the team at Metsamor to implement it into the daily recording system of the excavation. This resulted in a more thorough documentation throughout the season. Each excavated area was recorded in detail on a regular basis by the process discussed in this article. This led to reduced workload by eliminating the need for spending hours or days measuring and drawing in the field. The new approach can help archaeologists in decision-making on-the-go, while providing a comprehensive back-up of their work, which can be used even after the end of the season for interpretation and stratigraphy analysis.

Otto Bagi
otto.bagi@mail.com

REFERENCES

Callieri, M. et al. (2011). Documentation and interpretation of an archeological excavation: an experience with dense stereo reconstruction tools. In M. Dellepiane et al. (eds), *VAST 2011: the 12th International Symposium on Virtual Reality, Archaeology and Cultural Heritage* (pp. 33–40). Goslar: Eurographics Association

De Reu, J., De Clercq, W. et al. (2013). Orthophoto mapping and digital surface modeling for archaeological excavations. An image-based 3D modeling approach. In *Proceedings of the 2013 Digital Heritage International Congress (DigitalHeritage)* (pp. 205–208). Piscataway, NJ: Institute of Electrical and Electronics Engineers

De Reu, J., Plets, G. et al. (2013). Towards a three-dimensional cost-effective registration of the archaeological heritage. *Journal of Archaeological Science*, 40(2), 1108–1121

Doneus, 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 FCE CTU*, 6, 81–88

Grussenmeyer, P. et al. (2008). Comparison methods of terrestrial laser scanning, photogrammetry and tacheometry data for recording of cultural heritage buildings. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B5), 213–218

Kjellman, E. (2012). *From 2D to 3D – A photogrammetric revolution in archaeology?* (unpubl. MA thesis). University of Tromsø

Szeliski, R. (2010). *Computer vision: algorithms and applications*. New York: Springer

Verhoeven, G. (2011). Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photoscan. *Archaeological Prospection*, 18(1), 67–73

Verhoeven, G. et al. (2013). Undistorting the past: new techniques for orthorectification of archaeological aerial frame imagery. In C. Corsi, B. Slapšak, and F. Vermeulen (eds), *Good practice in archaeological diagnostics: Non-invasive survey of complex archaeological sites* (pp. 31–67). Cham: Springer International Publishing

Westoby, M.J. et al. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300–314