CREATION OF A 3D REFERENCE MODEL OF AN ARCHAEOLOGICAL SITE FROM A LARGE SET OF GROUND AND UAV IMAGES

E. Alby and T. Dellarovere
Université de Strasbourg, INSA Strasbourg, CNRS, ICube Laboratory UMR 7357, Photogrammetry and Geomatics Group, 67000 Strasbourg, France.
(emmanuel.alby, toma.dellarovere)@insa-strasbourg.fr

KEY WORDS: Photogrammetry, archaeology, automation, large data set, network computing, Cultural Heritage

ABSTRACT:

The archaeological site of the monastery of Saint Hilarion, in the Gaza Strip, concentrates important social and scientific issues. Its geopolitical context puts into tension the importance of setting up a pilot project for archaeological site management and the possibility of offering stable working conditions to many people. The impossibility for us to access the site requires us to delegate the acquisition work and leads to an increase in the complexity of data processing, while maintaining a requirement for the quality of photogrammetry products for archaeology. A large site implies a large amount of data to manage, delegated acquisition increases the volume even more and the impossibility of controlling the support points imposes the control of the data set based on the orientation results. These constraints imply important computational phases influencing the processing method.

1. INTRODUCTION

Photogrammetry is increasingly used by archaeologists. Its integration into the excavation process facilitates the continuity of the excavation of stratigraphic layers. The drawing phase of the survey no longer intervenes at this stage and makes the excavators' work more fluid. The photogrammetric survey introduces 2D data through the ortho-mosaic, in plan and cross-section, and 3D with the textured mesh. The importance that these data have taken on, gives them the role of reference representation and imposes centimetric precision, whereas remote management imposes a delegation of the acquisition and leads to methods integrating this contradiction.

The combination of a large site, ground and UAV data, numerous details, remote acquisition management and ground points that are impossible to verify reduces the room for manoeuvre. The search for automation of a maximum of processing is necessary.

2. STATE OF THE ART

2.1 White balance improvement

To be able to create quality textures that are close to reality, the issues of colour presentation and white balance need to be considered to standardize the different photographic campaigns. Indeed, these different campaigns produce a juxtaposition of images for one or more rooms of the site. To limit any artificial differences and demarcation, which would be detrimental to the aesthetic quality of the textures, radiometric harmonisation must be carried out.

One of the greatest challenges of colour in photography is how to acquire it efficiently and reliably to best reproduce reality as we perceive it with our eyes. The cameras used in this project use the standard Red Green Blue (sRGB) space, which over time has become the default operating mode for cameras. Although the colour space is the same, depending on the type of light source that illuminates an object you wish to acquire, the colour will not come out in the same way on the image. The aim of white balance is therefore to "maintain" the colour close to that of an observation in neutral lighting despite the presence of a coloured light source. In other words, we are trying to find a reference that will allow us to get closer to the human eye's perception of colour.

To correct the white balance in post-production, there are manual or automatic approaches. In view of the large number of images in our project (around 86,000), it is appropriate to look at the automatic approach that can be easily implemented. The article by (Zapryanov et al., 2012) provides a summary of these automatic approaches, which are called Automatic White Balance (AWB). In this paper, we are only interested in the choice between four automatic algorithms:

- Gray World Theory (GWT)
- Retinex Theory
- Gray World + Retinex Theory (GWR)
- Standard Deviation-weighted Gray World (SDWGW)

We studied different white balance correction algorithms using the article by (Zapryanov et al., 2012). The GWT and Retinex algorithms were found to be suitable for our situation. We therefore tested these two algorithms on samples of project images from different sectors. The GWT algorithm returns corrected images, which are duller and less bright, because the different channels are normalized by an average value, whereas the Retinex algorithm returns corrected images without loss of brightness because of the normalization according to the maximum of each channel as shown in figure 1.

Figure 1. result of the application of Retinex white balance.
Left: raw image without correction,
Right: same image with white balance applied.
2.2 Enhancement and optimisation of a dense cloud

Depending on the number of images used in a photogrammetric project, the size of the point cloud can vary greatly. For projects with a lot of images (> 20,000), both UAV and terrestrial, the cloud can have around 1 billion points, which is difficult to display for standard computers, but especially difficult to process for any type of computer. It is then necessary to think about an optimisation method because the density of these clouds is too high for our needs.

Cloud filtering has developed a lot in recent years, especially in the field of robotics and due to the increase in the size of the objects to be processed to have smaller datasets, to allow the acceleration of real-time calculations (Leal et al., 2017).

Whatever algorithm is used, it will consider either a physical or statistical characteristic to reduce the number of points. Most generally the reduction is done according to the neighbouring points except for the case of a random reduction (Han et al., 2017).

For our project and in view of the different existing approaches, we have chosen to use the topological approach

In this case, the reduction algorithms consider the position in space of the points in relation to each other when deciding whether to delete a point. These methods are not the most efficient for keeping the geometric information of the clouds, but they are still easily accessible and allow for simple density reduction without considering other more complex parameters.

The CloudCompare software uses a topology-based approach in its spatial filter that searches for each point within a user-defined perimeter to remove. Indeed, in addition to being programmable, which is relevant for our automatic processing chain, the software is open source, and we know how the algorithm works. Moreover, we are only trying to reduce the density. As the density is very high, a reduction based on the topology will cause little loss of geometric information in the cloud. This can then be seen as a noise reduction and not a true optimization of the point cloud.

2.3 Improvement and optimization of a 3D model

Models with sometimes hundreds of millions of faces cannot be easily managed by any computer. For the monastery of St. Hilarion, we consider this to be a model as acquired (Fuchs et al., 2004). This type of model represents the object as it is at the time of acquisition. The advantage of modelling a ruined site such as the monastery is that there are many flat surfaces, which can be significantly simplified.

When one wishes to reduce a 3D model, there are two possible ways, either one tries to simplify the model, which also reduces the file size, or one tries to reduce the weight of the model, to compress the representation (Cignoni et al., 1998). Here, we do not wish to compress the model, but to really reduce the number of faces, because they are overabundant. It should also be noted that the coordinate system, as well as the type of digital coding chosen, influence the final size of the 3D model.

Polygon model reduction algorithms can be grouped into three main categories (Garland and Heckbert, 1998):
- By grouping of vertices
- By reduction of the number of vertices
- By an iterative reduction of the number of edges

We have chosen to use the iterative one, as it allows a good compromise between speed and quality of the simplified model. We will use the rms-based reduction algorithm of (Cignoni et al., 1998) which gives good results and is available in Python, which is essential for the automation of the processing chain.

3. FRENCH-PALESTINIAN COLLABORATION

The Monastery of St. Hilarion is a major archaeological site located in the territory of the municipality of Nuseirat in the Gaza Strip in the Palestinian Territory. The site is approximately 1.5 Ha in size and its main occupation is dated to the 4th century (Alby et al., 2013).

As early as 2012, photogrammetry was chosen to simplify and speed up the representation process (Alby et al., 2015), but also to produce many useful documents for interpretation, whether 3D models, orthophotos, digital terrain models, sections, etc. To implement photogrammetry, it is necessary to be able to acquire the data and process it to achieve the expected results. The only constraint is that it is difficult to get equipment and personnel into the Gaza Strip.

It was agreed to work with the equipment available and accessible on site, to train architecture students from Gaza in photogrammetric techniques, to establish a stable Internet connection for sending the data, and to process them remotely.

Two training sessions on photogrammetry took place in 2020 and 2021. These trainings were given by French teachers and students. The objective was to make them aware of the importance of acquiring qualitative photographs, following a protocol whose strategy had been worked out previously. In addition to training in SLR cameras and the theoretical principles of photogrammetry, the architecture students were also trained in the Metashape software used for this project.

4. OPERATIONAL PROCESSING CHAIN OF THE PROJECT

The success of such a remote project relies on a good organisation between data acquisition and the creation of the different products. Of course, this organisation has its limits, but it is the best one considering the situation of the site and its constraints. The first step is the acquisition of images by the field team on the field. All the images are in JPEG format to have a compromise between weight and quality. Once the images have been acquired, they are renamed and arranged according to a standardised nomenclature and then sent to France via an online depository. Once received, they are saved. The images are then enhanced and added to different Metashape projects as required.

4.2 Data enhancement and sorting

One of the particularities of this archaeological project is that the acquisition of the images took place between 2012 and 2022 at different times of the project, but also with different cameras and light conditions. As a result, the white balance is not always correct, some images are blurred or of poor quality, etc. It is therefore necessary to correct these images as best as possible. Moreover, because of the amount of data, the manual approach is to be avoided, as it is too time-consuming.
4.2.1 White balance correction: The notion of white balance is essential for the consistency of the products resulting from the various computations. This result is far from uniform and needs to be improved. After this project-specific study, we decided to use the Retinex algorithm, which proved to be the most suitable for our situation, especially for scenes with natural light.

To apply this correction to all our 86,000 images, we developed a Python script. This script allows us to apply a correction to each image using both CPU and GPU at the same time without loss of metadata (EXIF). In addition, the process can correct approximately 3 images per second, making it a powerful script.

4.3 Production of the 3D reference model for the monastery site

In this section, we will see the processes implemented to more easily compute a very large dataset of nearly 86,000 images, the equivalent of 750 Gb of disk space, and to achieve a true operational computation strategy.

4.3.1 Implementation of a geo-referencing method for the model: The size of the monastery, as well as the different temporalities of the acquisitions, requires us to think about implementing a real geo-referencing system. Indeed, the georeferencing of the site, and consequently of all the products we create, has many advantages:
- it ensures the accuracy of the final data; we want to stay below 5 cm at the end of absolute orientation.
- it facilitates the process of aligning images, but also the merging of different projects.
- the existence of known targets in coordinates allows us to scale the projects and check their coherence.
- it allows us to have the products in the same coordinate system, which allows them to be reused in a GIS, but also to be overlaid.
- it also allows us to geo-reference the excavations, which allows them to be superimposed on more recent data.

To implement this geo-referencing, it was decided in 2020 to install 54 targets distributed between the different zones and to have them measured by a local survey team. All coordinates are in the local Palestine 1923 / Palestine Grid modified system.

During the first calculations, we realised that the coordinates of some of the targets did not match and had residuals greater than 10 cm. It was decided to carry out a new topographic campaign for all the targets. Despite our efforts to obtain control over the data produced, it was not possible to improve the acquisition method, and only a comparison with the old data set was possible. Indeed, among the 52 remeasured targets, only one has a 30 cm error on the x-axis after alignment processing. It will not be used for all the calculations. The new coordinates are more consistent and the deviations at the end of the photogrammetric calculations are smaller.

4.3.2 Strategy for aligning the site's images

General information: The monastery is a complex site both in terms of the elements to be modelled and the temporality and process of data acquisition. In addition, the monastery is subject to different types of work and includes different architectural elements. The soils of the site are varied and need to be treated with care. It is therefore necessary to consider all these particularities in the computing strategies that we implement, whether for local or global computing.

Regarding the number of images, no single computer at our disposal can efficiently calculate projects with more than 86,000 images. It is therefore necessary to find several strategies, or even one, that will allow us to achieve our goals and compute the entire site.

It is then possible to:
- calculate each part independently of the others, which makes it possible to use all the images, but only if the control points have been validated.
- create a single project by dividing the calculations by tile and merging them at the end of the calculations. Merging the calculated tiles will require a lot of resources, including RAM. A good computer will be needed to perform this operation.
- perform a network calculation to process large projects more easily. It is important that the nodes of the network are efficient in order to be able to process the different tasks which may require a lot of RAM depending on the number of images to be processed;

In our project, we tested most of these strategies to meet various objectives.

Over time, after several trials, we were able to develop a universal strategy for large-scale projects. In the rest of this section, we will look at different cases, including the universal case.

Development of a strategy for the calculation of several areas: The objective of this project is to align all 86,000 images to create a global 3D model of the site that would serve as a reference. Due to the large number of images, we developed and implemented a computational network proposed by Metashape.

Metashape must perform automatic preselections, which are very RAM-intensive tasks. When we tried to compute all 86,000 images in the project, the initialization of the alignment computation could not succeed due to lack of RAM, even though our most powerful machine had 128 Gb of RAM. Our estimate based on other smaller projects was that we needed about 430 Gb of RAM for the initialization, which was not available at the time of the test. We then modified the objective of this first large-scale network calculation and tried to calculate the whole site in two parts: one containing the images of the first two zones, i.e. nearly 56,000 images, and the other containing the images of three other areas, i.e. nearly 25,000 images. Once these two calculations were completed, we were able to align the images, calculate a model for each part and the corresponding orthophoto. The performance of the computing network enabled us to align 99% of the images of the two parts in 2 days and 1 hour.

This raises the question of the relevance of a network calculation. At first glance, following the two-part calculation of the site, it can be considered that the network calculation allows us to speed up the process and save time. This is a rather rational result, several computers calculating the same project at the same time go faster than one computer alone. Nevertheless, to have tangible comparison data, we calculated the same project on a single computer and via the network to see the time saving. This project consists of 26,800 images acquired by UAV alone. The computer alone aligned the images in 3 days and 12 hours compared to 20 hours for the computing network, which represents a time saving of nearly 76%! A computing network is therefore relevant for projects involving more than 10,000 images.

Development of a universal strategy for the calculation of a large project: The idea of developing a universal strategy for large-scale projects, i.e., projects where even a computing network reaches its limits, became clear when we realised the...
important need for RAM. Indeed, despite preliminary research, it is difficult to estimate the RAM requirements of a project under Metashape except by experimentation. It is therefore as we carried out the alignments that we realised the limits of the conventional technique, which consists of providing the software with many images without information on their external orientation parameters. In its operation, Metashape can consider in the calculation process the coordinates of the camera centres as well as their rotation parameters. As it is a new generation software, it is possible to give it images with and without its parameters, which allows to add new images over time, but also to consider those that could not be aligned in the past. Our strategy is based on two main principles and allows us to considerably reduce the need for RAM, the calculation time, and the alignment errors; these are the georeferencing of the project and the calculation by tile. The aim is to pre-compute small projects, the number of images of which depends on the computer power at our disposal, and then to reuse the approximated external orientation parameters as source parameters in the final project. The advantage of small projects is that they are easy to calculate and avoid alignment errors. The software will therefore have information to locate the images in relation to each other, which will help it to consider overlapping, make selections, but also to place the images correctly and avoid deviations due to large areas. This strategy therefore combines many advantages but requires quality georeferencing and therefore the presence of targets on the site. Moreover, this strategy can easily be implemented by a network.

To evaluate the performance of this strategy, we adapted our computational network so that it could align a project with the set of images at our disposal without using the approximated external orientation parameters. To do this, we modified one workstation to give it nearly 320 Gb of physical RAM with an additional 128 Gb of VRAM, bringing our available memory to 448 Gb. We estimated that we needed 430 Gb of RAM, so there is a margin of 18 Gb to make the calculation successful. The calculation was successful and did require less than 100 Gb of RAM in its initialisation phase. The alignment was good, as 98% of the images were aligned and few alignment errors were found. Once this first calculation was done, we repeated it by reusing the approximated external orientation parameters as source parameters and by adding new images that had never been aligned before. The result was very interesting, as the alignment was successful and required less than 100 Gb of RAM in its initialization phase.

The universal strategy we have developed is therefore relevant and allows for efficient calculation of large projects.

4.3.3 Improvement and correction of 3D models in Metashape: Now that we can align large projects efficiently, we can easily create 3D models using the depth maps. Indeed, the algorithm will first compute a dense point cloud, then keep only 10% of the pure points visible from at least five maps. Furthermore, if images are slightly misaligned, errors or multiple triangle layers may be found in the same location. Apart from the cleaning issues, there are also issues related to the size of the model and the number of faces. At the beginning of the project, we did not have any criteria for the number of faces per square meter. So, we gave the software a quality between low, medium and high. This black box operation provides us with models with several faces that is too high for our needs. The study of the different model reduction algorithms presented in paragraph 2.3, allows us to counter this effect and to reduce them efficiently.

Cleaning of a model: In general, cleaning a 3D model in Metashape is an action that can be easily performed using the various tools at our disposal, such as manual or automatic selection according to different criteria. The software is perfectly optimised to navigate within a model even if it has more than 100 million faces. This is a real advantage, as other less optimised software can make the experience more difficult. However, despite all its advantages, Metashape does not offer a “clipping box” that allows you to view and modify only a part of a model without influencing the rest. This lack can make the cleaning process longer, especially for details such as excavations.

In our case, we chose to apply a pyramidal cleaning method, starting with a global cleaning and ending with a more detailed local cleaning. This method is the most rational and avoids redundancies and therefore wastes time.

At the end of this three-step process, we have a corrected model that represents the desired state of the site and is ready for geometric improvement.

Optimisation using external and Metashape-specific functions: The cleaning of the model is only one step in obtaining a final model. Indeed, once the cleaning is done, it is necessary to optimise the model either for the number of faces or for the topology.

Metashape natively offers two very useful functions for topology correction: "Close Holes" and "Fix Topology". The first one allows to close the holes created during the cleaning phase; the software analyses the model and according to a value of percentage of size of the holes that we indicate, it closes them. The second function allows the topology of the model to be corrected, i.e. duplicate faces and edges to be removed, normals to be fixed and edges to be fixed. It is important to apply this function last to correct the final model.

Reducing the number of faces in the model can be done natively in Metashape. The problem with this approach is that we do not know the algorithms behind the function. We have therefore chosen to study different types of reduction algorithms seen in paragraph 2.3. We have therefore developed a Python script allowing us to apply this filter to a model using the MeshLab python library.

To determine the number of faces acceptable for our project, we carried out various tests in collaboration with the archaeologist in charge of the site. The objective was to determine the minimum number of faces necessary to ensure a good accuracy of the model for video and orthophotos. This requires:

- 6,000 faces per m2 for a model for normal use (to create textures, videos, orthophotos)
- 10,000 faces per m2 for a model representing excavations.

Of course, these rates can be used as soon as the 3D model is created to avoid the model reduction step. Indeed, we have the possibility to give Metashape the desired number of faces during the creation stage.

5. REUSE OF THE REFERENCE MODEL FOR THE PROJECT

We have seen above the process of creating a global 3D model of the site considering the constraints and the geographical location of the monastery. Now that we have a general, cleaned, and optimised model, we can see to what extent it is possible to reuse this model to improve the treatment process, the management of the site or the archaeological products.
5.1 Production of true orthomosaics

The archaeologist in charge of the site has relied from the beginning of the project on the creation of true orthomosaics to develop a new method of field acquisition, which would be faster and more cost-effective. The aim is to carry out excavation campaigns in a shorter period where the manual drawing of stratigraphic layers is replaced by image acquisition and photogrammetric calculations (Alby et al., 2015). Unlike manual drawings, orthomosaics allow for renderings with textures from real footage proceeding with good accuracy, usually around 2mm/pixel, as well as correction of various deformations and thus conservation of scale.

Once all the images have been geometrically corrected, the orthophoto will not be perfect and you will not be able to improve your data management by reducing the overlap between images or by classifying them geographically by room or area.

5.2 Image classification

In addition to enabling the creation of archaeological products such as orthophotos, the global 3D model of the site allows us to improve our data management by reducing the overlap between images or by classifying them geographically by room or area.

5.2.1 Reducing overlap

There are two ways of reducing the overlap and thus limiting the number of images for our project: the arbitrary method, without considering the calculated orientation of the images, and the a posteriori method, which considers the alignment performed. For our project, we chose to use the second approach.

To carry out this approach, we have developed a method based on the overlap reduction algorithm proposed by Metashape natively. In addition, we have developed a Python script that allows us to apply this method to all the photo files automatically.

Reduction values of around 60% can be achieved for some spaces. However, high values mean too much reduction and should never be reached. Despite the advantages of this algorithm, there are some limitations to its use. Overlay reduction is useful if all the images in the file are of the same quality and if you do not want to create complex or detailed models. Indeed, reducing the overlap always carries the risk of not considering the good images that can be used punctually. Another limitation is that we have no visual on the model created before reduction, because everything is done automatically by the python script which only renames the images according to the results.

5.2.2 Classification of images by room

Due to the organisation of our processing chain, all images are sorted by date of acquisition and by area in an approximate way. Indeed, the images are arranged in folders indicating one or more rooms. It is sometimes difficult to know exactly which room is visible on which image.

It is therefore interesting to find a method to classify the images geographically with precision. To do this, we only have aligned and georeferenced projects and a global 3D model at our disposal. The main problem in our case is to find a way to create a geometric link between images and locus. The method to create this link is only in 2D using an intersection of polygons.

To implement this method, we need polygons representing the boundary of each room, as well as polygons representing the footprint of each image on our global model. The fact that our projects are georeferenced makes it easy to make a geometric link between the different polygons, as the coordinates of their vertices will always be in the same system. Metashape provides a Python script to automatically create the footprint of each image aligned to the model as a polygon. The optimisation of the script consisted in implementing a multicore calculation adapted to our computer. The creation of the boundary polygons of each space is done manually, as they are set up by the archaeologist and

---

Figure 2. 3D reference model of the site where all the images from the 2020-2022 campaign were used.
depend on archaeological criteria. Indeed, there is no information
on the model that allows the automatic creation of these
polygons. It was therefore necessary to create each of the 162
rooms boundary polygons by hand, taking care to ensure that
each polygon slightly overlaps its neighbours. The use of the
Python package shapely allows us to use the boolean function
intersects() which tests the intersection between two polygons.
We then test for each of the boundary polygons, the interaction
with the set of camera footprint polygons. The output of the script
is a Python dictionary with the name of the locus as a key and a
list of all images that see the locus as a path to the images as a
value. The entire process can be seen in Figure 3.

5.3 Integration of excavation projects
It is sometimes necessary to align excavation campaigns to create
archaeological products showing different states. Most of the
photogrammetric survey campaigns related to archaeology were
carried out between 2012 and 2020. The methodology used at that
time did not incorporate the use of targets. In this context, a solution
must be found to align these campaigns, but above all to georeference
them so that they can be superimposed on other site conditions.

While making an inventory of the data at our disposal, we
realised that with the creation of our general model, we know the
approximate external orientation parameters of the images, i.e.
their position and orientation in space. We therefore decided to
add to the calculations of the excavation projects recent or older
ingages, corresponding to the excavation area, and whose
approximate position we know. After the alignment, it will be
sufficient to add the approximate coordinates as source
coordinates and to use the camera centres and not the targets to
georeference the projects. After several trials, the alignments
work and the reprojection errors on the control points, which are
the targets, are between 1 and 5 cm. It can therefore be concluded
that our method of integrating excavation projects is effective and
relatively accurate. Furthermore, to ensure optimal operation, it
is advisable to use mainly terrestrial images showing the
boundary walls of the locations. Indeed, between the excavation
projects and the recent acquisitions, the boundary walls have
changed very little, which allows us to have a temporal link
facilitating the alignment.

Figure 3. Schematic diagram of the working principle of the
room classification method.

Figure 4. Example of the intersection of the image footprint
(green) with the locus boundary (white).

Figure 5. Working principle of the method for the integration of
the excavation projects.
The 3D reference model is a real basis for improving data management or creating products such as true orthomosaics. Indeed, by studying the 3D model, it is possible to reduce the overlap between images, to classify images by locus or to georeference excavation projects. Here, the reuse of the 3D model is limited to certain needs, but all possibilities can be envisaged, because the model is a real common thread between the different data that we have or that we can produce.

6. CONCLUSION

The management of the representation of a large archaeological site at a distance requires a particular method. The 3D reference model of the monastery of Saint-Hilarion allows us to make the different types of photogrammetric data consistent and to validate the georeferencing acquired by delegation. The model itself allows the generation of classic photogrammetric representations but also allows the referencing of past and future projects of the site and thus the coherence of all data. The next step, beyond the referencing of all photogrammetric projects, is the creation of a dedicated archaeological information system. The method set up will be extended to many other projects on the same territory.

REFERENCES


Xian-Feng HAN, Jesse S. JIN, Ming-Jie WANG, Wei JIANG, Lei GAO et Liping XIAO : A review of algorithms for filtering the 3D point cloud. Signal Processing : Image Communication, 57:103–112, septembre 2017. 16
