FISHEYE PHOTOGRAMMETRY TO GENERATE LOW-COST DTMS

L. Perfetti 1,*, F. Fassi 1, C. Rossi 1

1 Politecnico di Milano, Dept. of Architecture, Built environment and Construction engineering (ABC), Italy
Email: <luca.perfetti><francesco.fassi><corinna.rossi>@polimi.it

Commission II

KEY WORDS: Fisheye, photogrammetry, low-cost, DTM, DSM, archaeology

ABSTRACT:

In the archaeological practice, Digital Terrain Models (DTMs) and Digital Surface Models (DSMs) may be used to represent spatial information about the site by conveying information such as differences in levels, morphology of the terrain and movements of volumes during the excavation. Nowadays DTMs and DSMs can be easily obtained by image-based matching using low altitude aerial dataset acquired from a digital camera by means of a lifting device. In recent years, the spread of commercial multi-rotor unmanned aerial vehicles and their decreasing cost made low-altitude aerial photography even easier than before, where balloons, kites and telescopic masts would have been used instead. However, the use of drones is often forbidden by law, especially in the archaeological areas, and therefore a more traditional approach must to be adopted instead.

This paper presents two different approaches adopted on the field to acquire the DTM of an archaeological excavation: the use of a pole held by a chest harness to lift a camera up to 3.5m height fitted with a 20mm wide angle lens; and a second solution that exploits ground-based fisheye photogrammetry. In general, an image network acquired from ground level is challenging due to: i) the poor coverage that can be obtained on the ground, ii) the large number of images that are required to cover large areas and consequently iii) the longer elaboration time that is required to process the data. The fisheye approach, however, proved to be more effective thanks to the more robust image network resulting both from the wider field of view and from the possibility to handle large datasets by downsampling the images and still retrieving strong key points. The main difference with the first system is that the monotonous images acquired by the 20mm lens, very plain in texture, require working at full resolution in order to distinguish valid features in the sand.

The final product of the tests carried out along this line in 2019 at Saqqara (Egypt) is a comprehensive DSM of the entire archaeological site with an accuracy of ~3cm.

1. INTRODUCTION

1.1 3D modelling of excavation areas

Digital Terrain Models and Digital Surface Models describe the 3D morphology of an area considering, respectively, the terrain only and the terrain plus all the structures and vegetation that cover it. They are widespread products that can be used to digitally represent a portion of a territory and to add 3D information to cartographic and Geographic Information System (GIS) environments. They are mostly used for small-scale territorial representations of large areas, but also, at larger scale, in the field of cultural heritage. DTMs and DSMs are of great importance in archaeology as well to record spatial information (Campana et al., 2007; Verhoeven et al., 2012). For example, in the archaeological practice, DTMs can be useful to understand differences in levels, to evaluate and study volumes excavated in the past, and to monitor current excavation activities.

In order to effectively document an archaeological site both three-dimensionally and photographically, several solutions have been developed during the last decades. The aim was to achieve the so called Low-Altitude Aerial Photography (LAAP) that could grant archaeologist high resolution aerial images characterised by a wide Field Of View (FOV), to complement the ground-based documentation. Lightweight lifting devices like balloons, kites, masts, poles etc., each with its own advantages and disadvantages, became common in the field practice to lift a camera into the air (Verhoeven et al., 2009; Chiabrando et al., 2010; Mozas-Calvache et al., 2012).

The advent of LIDAR sensor and the spread of Multi View Stereo (MVS) algorithms made it also easier to record high-density/small-sampling data to produce accurate 3D models. The most significant leap forward came from the spread of commercial multi-rotor Unmanned Aerial Vehicles (UAVs) that, due to their increasing accessibility, allowed low-cost aerial photogrammetry based on MVS algorithms to become common practice, by making the production of large-scale DTMs faster, more accessible, and more affordable (Rinaudo et al., 2012; Nex et al., 2014; Ballarini et al., 2015). In this scenario, affordable solutions grew exponentially and experimentation on the metric reliability of low-cost sensors and devices increased as well (Pagliari et al., 2018; Roncetti et al., 2018).

However, in several cases UAVs and other automatic modern instruments (including Terrestrial Laser Scanners and Global Navigation Satellite Systems) cannot be used on the field, due either to operational reasons (transportability, lack of electricity, environmental issues) as well as legal constrains; these issues often affect archaeological areas. In these cases, other technical and methodological solutions must be identified and adopted (Fassi et al., 2015). Alternative solutions must be based on tools that are allowed, such as the more traditional unmanned lifting devices (balloons or kites, although strongly dependent on wind condition), or masts that can be easily operated from the ground (Mozas-Calvache et al., 2019).

* Corresponding author
1.2 DMTs from fisheye terrestrial photogrammetry

This paper presents a low-cost method to extract DMTs of archaeological sites by employing close-range terrestrial photogrammetry, opposed to classic, low-altitude photogrammetry from masts or poles. The case-study was offered by the Dutch-Italian mission to Saqqara (Egypt), where drones are not allowed. Although there are several ways to lift the camera high over the ground to acquire the surface model of the terrain, the experiments carried out in 2019 suggest that a hand-held ground-based image network represents a fast and practical solution, that can be compared to the pole-aided method.

2. THE CASE STUDY AND THE REQUIREMENTS

2.1 The Leiden-Turin excavation concession area

The tests presented in this paper were conducted within the framework of the joint Dutch-Italian archaeological mission to Saqqara of Museo Egizio, Torino and Rijksmuseum van Oudheden, Leiden.

The mission’s concession measures ~14000 square metres, in shape it is almost rectangular and measure ~160m on the long side (east-west) and ~100m on the short side (north-south). The area includes the remains of several New Kingdom temple-tombs, including those of Horemheb, Maya, Tia, Meryneith and Puthemwia, excavated from the 70’s onwards (Martin, 1989, 1991, 1997, 2001; Raven 1991, 2005; Raven and van Walsum 2014) and now open to the public after a significant intervention of consolidation and reconstruction (Warner, 2009). The area currently under investigation is located in the north-west quadrant of the concession. All around the excavated area, the terrain is covered by the spoil heaps dumped by previous excavations.

From 2018, in the collaboration with the 3D Survey Group of Politecnico di Milano, a new survey of the area was started, with the aim of producing a first digital documentation of the concession area. The work proceeded along three parallel lines: i) the daily recording of the stratigraphy that was removed during the excavation, ii) the 3D survey of the already excavated and consolidated tombs, and iii) the 3D survey of the terrain around the discovered structures (Del Vesco et al., 2018). Two seasons (2018 and 2019) have been completed at the time of the publication of this article.

2.2 A unified high-resolution 3D model

This investigation started with the aim of generating the DTM of the mission’s concession (Figure 1) in order to complete the DSM, together with the detailed survey of the structures that punctuate the area (Figure 6). In turn, the aim of the unified DSM is to describe the whole area in 3D and allow a study of the excavated levels and the amount of dumped material, as well as an analysis of levels to be excavated in the future.

The resolution for such model can be about 10cm, a value which is relatively low if compared to the resolution of the models of the tombs that are being prepared in parallel (0.5/10mm); therefore, such an approximate value is compatible with pole-aided or ground based photogrammetry.

3. THE SURVEY METHODS

A terrestrial photogrammetric survey to generate DMTs can be challenging because of the many issues affecting the acquisition of data, such as: the extremely close capturing distance, the great number of photos that are required, the complex topographic geometry and the mandatory use of a sufficient number of Ground Control Points (GCPs) / Check Points (CPs) to constrain the acquisition.

Some initial tests were carried out by means of a chest-mounted pole used to lift a standard Digital Single Lens Reflex (DSLR) camera. A pole serves the purpose of increasing the capturing distance, i.e., the distance between the camera and the object, as well as to increase the image coverage on the ground, thus reducing the number of images required to cover the whole area with sufficient overlap.

As an alternative to the more conventional pole-aided photogrammetry, fisheye photogrammetry was used in a second survey phase. Fisheye lenses convey a wider FOV onto the sensor compared to rectilinear lenses, resulting in a larger Ground Sampling Distance (GSD), thus generating lower resolution images.

Table 1 reports the data obtained from the two methods. The number of photos that are required, the complex topographic geometry and the mandatory use of a sufficient number of Ground Control Points (GCPs) / Check Points (CPs) to constrain the acquisition.

Table 1 reports an estimate of the number of images that would be necessary to cover the same area covered by the fisheye method.

Figure 1. Image network carried out to survey the DTM on the concession of the Dutch-Italian mission to Saqqara.
The POI used was a chest-mounted pole, and the speed of the acquisition was faster by using downsampled images, but still retrieving good features (Figure 4). The total amount of images acquired to cover the entire area were 6200 and were taken from a height of ~2 metres. The whole survey was conducted by a single operator by walking in the concession area. The camera was oriented at approximately 45° pointing forward and held up with the arm extended. On this occasion, the operator walked straight ahead, and not sideways, in parallel stripes few metres apart (Figure 1). The resulting overlap between consecutive images exceeded 90%, and the GSD, although hard to estimate, was around 3-4mm at the centre of the frame and quickly increased towards the border of the frame as a result of both the increasing distance from the object and the fisheye projection function (Perfetti et al., 2017); overall, the average GSD can be estimated at one order of magnitude greater than the rectilinear configuration: 1.5cm. 

The image network consisted of a parallel linear acquisition that covered the entire terrain from immediately outside the concession area up to the architectural structures. The direction of the camera changed throughout the survey roughly following an anticlockwise direction revolving around the core of the area. To strengthen the network even more, images have been added along ‘arches’ around the outer corner of the reconstructed tombs. The survey was divided in four areas and it was carried out in four different days during the same period of the day, to guarantee similar conditions of illumination and shadows throughout the dataset. Overall the entire survey required eight hours on the field. The image capture was easy, versatile and fast, if compared with the clumsier acquisition carried out with the use of the chest-mounted pole, and the speed of the acquisition was limited just by the camera buffer.

### 3.2 Fisheye ground-level photogrammetry

After the first unsuccessful test, the acquisition was repeated during the 2019 season. This time a different strategy was followed. The goal was to surpass the limitation of the previous approach by, first of all, increasing the image overlap and the overall robustness of the network, and secondly, by providing lower level key points (2D features expressed in image coordinates that describe relevant points/features) used to align the images. Both goals can be achieved by employing a fisheye lens instead of a rectilinear one. Indeed, the FOV of the fisheye lens provides a wider coverage, and thanks to this increased coverage also the complexity of the images increases. Figure 3 on the right shows an image acquired during the survey; parts of the architectural structures are depicted together with the terrain, thus providing stronger key points compared to the 20mm lens image. When implementing this method, we decided not to use the pole in order to conduct a more rapid and agile acquisition, and to avoid including the operator in every image. Moreover, due to the already significant FOV, the additional gain would have been minimal. The low-level key points also allow the dataset to be process faster by using downsampled images, but still retrieving good features (Figure 4).

The first tests were conducted during the 2018 season (Del Vesco et al., 2019). A photogrammetric network was acquired thanks to a DSLR camera (Canon 5DmkIII) coupled with a 20mm rectilinear lens, from a height of approximately 3/3.5 meters from the surface. The GSD of the configuration was on average around 1.5/2mm. The operator carrying the camera followed a continuous walk in the concession around the structures to be surveyed, while a second operator acquired the images by means of a remote controller. The camera was tilted forward at an angle of about 45° and the operator moved sideways, thus guaranteeing sufficient overlap between the images on the horizontal direction. The images from the height of the pole could frame only small portions of the ground and the only valid features that could be used as key points for the image matching were the small details of the terrain, pottery fragments and stones (Figure 3 on the left). This first approach failed due to the complexity of the terrain and the poor and repetitive texture of the images. This method would require a longer pole or a higher number of images to increase the overlap, especially between different parallel ‘rows’ of sideways walks. However, the pole was already used at the maximum extension and a higher number of images would have made the method impractical. The reason is that the dataset would request to be processed at full resolution in order to fully exploit the fine details of the sandy terrain.

### Table 1. Specs of the two methods:

<table>
<thead>
<tr>
<th></th>
<th>Wide-angle</th>
<th>Fisheye</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>5760 X 3840</td>
<td>2900 X 2900</td>
</tr>
<tr>
<td>Lens</td>
<td>20mm</td>
<td>8mm fisheye</td>
</tr>
<tr>
<td>N° Images</td>
<td>4500 (2)</td>
<td>6200</td>
</tr>
<tr>
<td>FOV diagonal</td>
<td>94.5°</td>
<td>~150° (3)</td>
</tr>
<tr>
<td>GSD</td>
<td>~1.5/2mm</td>
<td>~4mm</td>
</tr>
<tr>
<td>Height</td>
<td>~3/3.5m</td>
<td>~2m</td>
</tr>
<tr>
<td>Overlap</td>
<td>~80%</td>
<td>~90%</td>
</tr>
</tbody>
</table>

3.1 Wide-angle lens and pole

The images from the height of the pole could frame only small portions of the ground and the only valid features that could be used as key points for the image matching were the small details of the terrain, pottery fragments and stones (Figure 3 on the left). This first approach failed due to the complexity of the terrain and the poor and repetitive texture of the images. This method would require a longer pole or a higher number of images to increase the overlap, especially between different parallel ‘rows’ of sideways walks. However, the pole was already used at the maximum extension and a higher number of images would have made the method impractical. The reason is that the dataset would request to be processed at full resolution in order to fully exploit the fine details of the sandy terrain.
The obtained datasets were processed in Agisoft Metashape at low-resolution. An initial calibration of the 8mm fisheye lens mounted on the Canon 5DmkIII was retrieved from previous works conducted during the mission, and subsequently adjusted on each dataset to fit the camera configuration. The latter changed every day for every block of images, due to the continuous switching of lenses that was essential to follow the excavation activities. To guarantee a good quality of the self-calibration, the operator continuously rotated the camera along the optical axis during the acquisition process (Figure 5). An APS-C lens cap was used on the lens to preventively mask the outer border of the fisheye projection, in order to discard the lower resolution portion of the frame (Perfetti et al., 2017), as well as the presence of the sun and of the operator’s feet or body. The shadow of the operator, however, is present in most of the images (Figure 5). The maximum FOV obtained with the masked images is around 150°.

4. EVALUATION OF THE RESULTS

4.1 Accuracy check

To improve and to check the accuracy of the obtained reconstruction, a set of GCPs and of CPs were used. The targets were manually picked in some images of the photogrammetric block using both targets printed on paper and fixed over the tombs’ walls, and relevant architectural features inside the concession and further away. A subset of the total amount of target used in the project were measured with a total station from the vertex of a network distributed all around the concession. The Root Mean Square Error (RMSE) on the CPs was around 3cm for all the four blocks.

4.2 DTM processing

The four oriented blocks where subsequently further divided in small area to process the final point cloud. The MVS algorithm of Agisoft Metashape used the images downsampled by a factor of four (low resolution). The resulting point cloud have been merged, cleaned, filtered and downsampling further to get rid of the fine detail of the terrain and to preserve only the overall levels. The point density after the subsampling averaged 2cm. Acquisition on the field and data processing to construct the DTM required 8 and ~80 hours respectively. Prior to the generation of the DSM, the downsampled point cloud of the terrain was merged with the point cloud of the tombs that were surveyed during the 2018 and 2019 seasons (Del Vesco et al., 2018). These point clouds had been also downsampled to a resolution of 1cm. The resulting point cloud that describes the DSM was used to generate the final mesh model (Figure 6).
5. DSM AND FUTURE USE

The resulting digital surface model allows to combine together different specific and punctual investigations (i.e. the reconstruction of small elements, of the toms and of the current excavation activities) as well as to place them in a broader picture that allows for more sophisticated and complex queries. Some of the possible products that can be extracted from the obtained DSM are listed in the following paragraphs.

5.1.1 Elevation map: the elevation map (Figure 7) was extracted from the mesh model of the concession at a resolution of 2cm; contour lines were extracted as well at a pace of 20cm.

5.1.2 Orthophoto: The orthomosaic of the whole area can also be retrieved from the surface model and the oriented images. Figure 8 shows a small portion of this area, the backfilled tomb of Ry. The orthomosaic shows a clear distortion, visible in the presence of sharp edges due to the DSM low resolution. In order to obtain the final product, a substantial amount of work was required to remove the operator and its shadow from the ortho projection.

5.1.3 Cross-sections: from the DSM and the original point clouds it is possible to extract cross sections of the entire concession. Moreover, repeating the survey day-by-day to record the excavation progress allows the opportunity to compare the movements of volumes (Figure 9).

DTMs and DSMs can find different applications in archaeology: moving from the material to the digital realm, archaeologists can link information and data in a more effective and productive way in comparison with the past (e.g. Rechichi et al., 2016). The complete DSM of the area, that combines terrain and remains (Figure 6) can be used to better understand the history of the site and the changes that took place both in ancient and recent times. Moreover, the possibility to combine the survey of the surface with that of the subterranean portions of the tombs, can provide an unprecedented insight into the stratified and complex evolution of the site (Rossi, 2019).
6. CONCLUSIONS

The results of this test demonstrated the possibility to generate high-resolution and accurate DTMs from the ground by employing fisheye lenses providing large FOV and coverage, whenever the use of UAVs is forbidden by law.

The described method proved to be more effective than the use of a rectilinear lens (20mm) mounted on a short pole (3.5m). It allowed us to process a large dataset using images which were downsampled by a factor of four; all this took place in a relatively short time, thanks to the robustness of the image network and of the key points extracted. The main advantage of a ground-based approach is the versality of the system as well as its affordability. In our case, while one operator was conducting the acquisition, a second operator could continue to follow and record the excavation progress without causing delays.

ACKNOWLEDGMENTS

We wish to thank Dr Christian Greco and Dr Lara Weiss, Co-Directors of the Dutch-Italian mission to Saqqara of Museo Egizio, Torino, and Rijksmuseum van Oudheden, Leiden for allowing us to experiment advanced survey techniques on their excavation. Thanks also to Dr Paolo Del Vesco, Field-Director, for his support during the survey operations.

This article is an outcome of the project L.I.F.E., that received funding from the European Research Council (ERC) under the Horizon 2020 research and innovation programme (grant agreement No. 681673).

REFERENCES


