Innovative Strategies and Use of UAVs to Survey and Monitor Archaeological Sites in Conflict/Post Conflict Zones. The Case Study of the Fortified Citadel of Shahr-i Zohak in Bamiyan (Afghanistan)

Daoud Bouledroua [1], Brendan Cassar [2]

[1] aeGIS Heritage Management, 75004 Paris, France – d.bouledroua@aegis-heritage.com
[2] Head of Culture, UNESCO Kabul field office, Afghanistan– b.cassar@unesco.org

Keywords: Architectural Heritage at Risk, 3D Reconstruction, UAV Photogrammetry, Flight Path Planning, UNESCO World Heritage Site, Damage Assessment.

Abstract

Although unmanned aerial vehicle (UAV) mapping and photogrammetry have become common and relatively accessible for surveying and mapping cultural heritage sites, conducting surveys to model sites in conflict/post conflict zones remains challenging. This is particularly true for sites in a country like Afghanistan, where limited accessibility, the presence of Unexploded ordnance (UXOs), portability of the equipment, cost efficiency, as well as absence of data connectivity and Ground Control Point establishment pose major challenges. In this paper, we discuss the adopted strategy and implemented methodology to create a 3D model from both inside and outside of a section of the fortified citadel of Shahr-e Zohak which is part of the UNESCO World Heritage Property of the Cultural Landscape and Archaeological Remains of the Bamiyan Valley in Afghanistan. In particular, we examine in this paper what acquisition strategy was set for this site by going through the reasoning for selecting specific equipment and drones, the flight parameters, the camera settings, as well as how we prepared the dataset at the flight planning stage to allow merging GPS referenced data from the external flights of the UAV with non-GPS referenced data from the flights inside the domes and built structures. Succinctly, we go through the modelling strategy and parameters that have generated optimal results using both Agisoft Metashape and Bentley Itwin Capture.

Our results show that using Skydio’s X10 and S2 drones and setting a low Ground to surface distance (between 1 and 5 meters) and high overlap (75%-95%) allowed us to achieve 3D models with an average accuracy of 1 millimetre per pixel for a 120m long and 30m wide section of the fortified citadel of Shahr-i Zohak. These results also show that it is indeed possible to use UAV based photogrammetry to generate 3D models that can be used for damage assessments which is particularly useful in areas where it is difficult or impossible to bring international experts or institutions to conduct this work on site. Finally, this research highlights the capabilities as well as limitations of this method and provide practical guidelines for future works in comparably challenging environments.

1. Introduction

The fortified citadel of Shahr-i Zohak is situated about 20km east of Bamiyan and is believed to have been founded in the 6th-7th century A.D. Built at a strategic location overlooking both the Kalu and Bamiyan rivers it commands the access to Bamiyan and is part of the UNESCO World Heritage Property of the Cultural Landscape and Archaeological Remains of the Bamiyan Valley in Afghanistan. Within the context of UNESCO’s ongoing project to safeguard and protect the World Heritage Site of Bamiyan, the decision was taken to create a 3D model of the citadel at a quality that would allow damage assessments and the planning of conservation interventions. This intervention is particularly timely as severe climate in the central Highlands of Afghanistan as well as the years of war have resulted in considerable deterioration of the state of conservation of the architectural and archaeological remains of Shahr-I Zohak. The site is composed of different earthen architectural defence structures leading to the summit of the hill to the south of the site at 200m above the ground. The structural remains of the citadel are situated on a triangular plateau of an area of approximately 1 hectare in the north of the hill at 50 to 70 m above the ground level. Current intervention focuses on this triangular plateau and therefore the focus of the survey was to be the set of domes that compose the remains of the citadel there (see circular rooms A to G visible on Figure 1). Preliminary research has shown that no major archaeological intervention has been conducted at the site and that the scientific research was sparse (the exception being P.H.B. Baker and F.R. Allchin, 1991 which includes a nomenclature of names and sites that was used for the survey).

Figure 1. Plan of Shahr-I Zohak (Baker, P.H.B., Allchin F.R., 1991, p. 48).

Being situated in the central highlands of Afghanistan, access to the site is difficult and there is no access to internet anywhere near the site. In addition to this, although some demining has been undertaken there is still a high risk of presence of unexploded ordnance (UXOs) and it is therefore dangerous to walk around the site outside of clear paths. Another challenge for surveying work in the area is that there has never been a national topographic survey that could have provided GPS corrections. Due to the steepness of valleys and height of surrounding mountains GPS access is also often degraded. It became therefore quickly clear that GNSS augmentation would be necessary and that it would not only be required to work with ground control points but also necessary to create a network of benchmarks set in concrete around the site. At the same time, the security situation at the site does not allow for prolonged stays of international experts and therefore the question of portability of the equipment was yet another impediment.

What would be the ideal surveying method in the context of all the above-mentioned limitations that would allow the creation of a 3D model that would be good enough to conduct damage assessments? In other words, what method would allow to create in this context a 3D model that would allow experts to conduct damage assessments and to analyse and measure cracks for example that are only millimetres wide at the same time as providing a basis that allows for planning of conservation interventions at the site?

Based on the survey conducted in November 2024 in Shahr-I Zohak, this paper proposes a comprehensive methodology and approach with a guideline about how to achieve the above set results for cultural heritage sites located in conflict or post-conflict zones. Proposing such a methodology is not only a good thing in itself but a requirement as laid out by the 2006 London Charter for the Computer-based Visualisation of Cultural Heritage, which highlights the importance to transparently and clearly document the modelling process ('paradata') and make it available so that others are able to understand and evaluate the actual quality of the underlying empirical analysis.

2. Acquisition strategy

Figure 2. Image of the south side of the domed structure

The mountainous and inaccessible terrain has immediately made us tend towards using unmanned aerial vehicles (UAV) for the mapping process. As UAV based photogrammetry has become increasingly accessible, popular and accurate (see for example Küng Olivier et. Al, 2011), we set out to find a drone that would be able to fly indoors and outdoors, detect obstacles easily, all the while being resistant against the harsh climate and conditions of Afghanistan (dust, strong winds, intense sun, and potential rain). The US based company Skydio offers two drones that correspond to these requirements: the larger X10 (equipped with a 20mm equivalent 93° field of view lens and a 50mp sensor) that was used for mapping the outside of the domes, walls and towers, as well as the smaller S2 (equipped with a 20mm lens and a 12mp sensor) that was used for mapping the inside of these structures. These UAVs were chosen as they have the capacity to fly in GPS denied environments (and can therefore be used to fly inside structures and buildings), but also and probably more importantly because they have the capacity to avoid obstacles and to calculate automatically the optimized flight path for the mapping, which allows to have constant coverage on all surfaces of the area that was to be surveyed. The specific area mapped in the context of this case study consist of a series of 6 circular rooms arranged in pairs and an enclosed rectangular room connected by a defensive wall and due to the complexity of the structure and requirement to achieve the above-mentioned quality it was necessary to capture the data in different segments.

Figure 3. Image of the north side of the domed structure

As can be seen in figure 2, and figure 3, the structure that we set out to scan posed different challenges both in terms of access, as well as in terms of planning the drone flights outside and inside the structure. The south side is very steep and difficult to walk on. Also, as the site was used at least throughout the last 50 years for diverse military purposes, the whole area is potentially contaminated by UXOs. Therefore, walking outside clear paths is not recommended.

With regard to flying drones in the area, the first challenge was the very bright and harsh mid-day sunlight (that is further exacerbated by the very dry high-altitude climate of the central highlands of Afghanistan). The second problem arose from the easy loss of connection to the Radio Controller when flying over the structure from the north to the south side (the gentler terrain on the north side meant that we would always have to keep the landing pad for the drone on this side of the structure). The third and probably most critical problem was linked to the way Skydio drones autonomously analyse the area that is to be modelled. Indeed, if the area is too big or complex, the drone will run out of battery in scouting phase of the scan. If this happens the whole setting of parameters must be started again, and the actual scan cannot begin. Therefore, structures that are too big cannot today be scanned by Skydio drones in one session and should be separated into smaller areas (an area of approximately half of an acre, which is approximately the size of the structure that was scanned had to be done in three session over three days).

With regard to the inside of the structure, the main challenges consisted of the narrow pathways, the lack of light in some parts and most importantly the dust. Room E was particularly challenging in this regard, as it has no real access to the outside except for a small opening in the roof and a small opening to the south side, which leaves this room with nearly no light. This also means that flying drones in these spaces would generate significant amounts of dust, interfere with the autonomous flight engine of the Skydio 2 and therefore render flying the drone in this area unsafe (therefore this room had to be scanned manually with our DSLR (a Sony A7III with a 35mm prime lens).

All this meant that the whole scanning procedure had to be carefully planned and adjusted onsite. The acquisition of the area was conducted over three days (with different lighting conditions). To mitigate against this fact, we decided to take the images of the outside at the same time each day (in the afternoon) and the images from the inside in the morning.

The acquisition of the outside of the structure was scanned in three sessions with the Skydio X10. The first area extended roughly from tower 17 to circular room F, the second from circular room B to room A, and the last from circular room C to D. To provide extra support for the alignment of the outside area four ground control points (GCPs) were set around the structure that would be visible from different angles and locations. These points were measured with a GNSS rover (for this survey we used the Emlid Reach RS3). Extra care was given here to use the same geodetic system for the GNSS and drones.

For the outside structure we aimed at a Ground to surface distance (GSD) of 5m and a general horizontal and vertical overlap of 75% between the images taken. The camera settings for the Skydio X10 were set to the recommended settings by Skydio for 3D modelling (1/4 of the resolution for each image and HDR turned on).

The inside structures of the domes were scanned in six main parts as well as three smaller ones for the passageways between the circular rooms in order to help with the alignment at a later stage. For the acquisition of the inside of the structure in nine parts (one capture for each circular room and one for each passageway) we aimed at a GSD of 1m and 85 to 95% horizontal and vertical overlap (as there are several decorative elements present within the domes it was necessary to obtain a enhanced quality from the inside of the structures).

As the inside of the structure is by default GPS denied and knowing that the most challenging aspect for modelling this site would be the alignment between the inside and outside segments of the structure, we made sure that every scan conducted inside a room or passageway also included images of the outside of the structure.

To facilitate the alignment of the images taken inside the structures with the outside we always made sure to start from the outside by taking several images from different angles from the entrance of a specific dome or tower and enter the structure then and start the actual acquisition of the inside of the structure.

Every scan of a dome was always scanned with the same drone and since we used the same brand for both drones that were used, we were able to make use of Skydio’s 3D scan application. After selecting via AR on the Radio Controller the area to be scanned, the drone then autonomously flies the most efficient and effective data capture mission and calculates the best flight paths to ensure every detail is captured homogeneously. Using this software, the drone will fly first horizontally around the area to be scanned and then in a crisscross pattern over the area in vertical lines. Figure X below shows how the scene appears in AR on Skydio’s Radio Controller. After each completed scan, we then took images with the drones of areas where the onboard AR engine showed that the desired overlap was not achieved.

By using this onboard autonomous flight planning and scanning software as well as by taking images manually with the drone of areas that didn’t have the required overlap, we were indeed able to achieve the set-out GSD and overlap ratios mentioned earlier.

Figure 4. Example of the horizontal flightpath autonomously generated for a tower structure with Skydio 3D Scan

3. Modelling and processing the data

We conducted initial trials with 4 different software applications, PIX4Dmatic, Reality Capture, Agisoft Metashape and Bentley Itwin Capture. For this early test we only used the images from the outside of the structure and tested the alignment with each software to see howe they would behave on the “simplest” aspect of the alignment. The outside scan is composed of around 6000 pictures. For the whole processing two machines were used: a MacBook Pro M4 Ultra with 48GB of ram; as well as an Intel I9-14900 Laptop with 192GB of ram.

PIX4Dmatic

The alignment done on the Macbook Pro with Pix4Dmatic took a very long time (nearly 20h) and the alignment that was generated with this software had a lot of noise (as can be seen in figure 5 below). The chosen parameters were the following: Pipeline: Standard; Images scale 1/1; Keypoints Auto; Internals confidence Low, Automatic ITPs activated; Compute relative confidence activated; Vertex conversion deactivated.

https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-171-2025

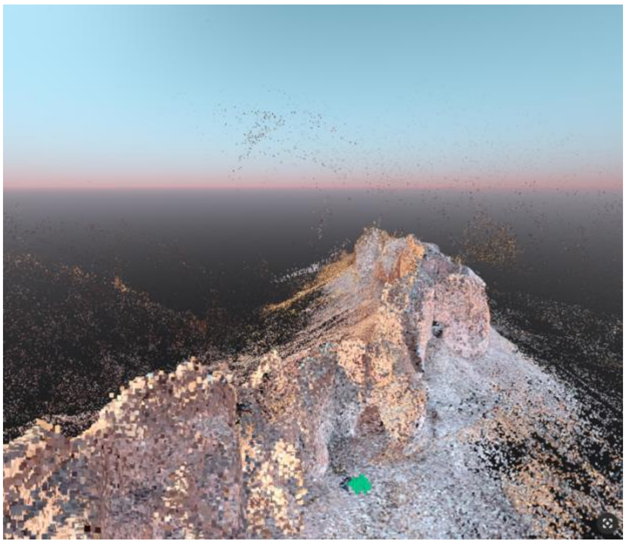

Figure 5. Alignment of the structure in Pix4Dmatic

Reality Capture

Reality Capture, although being much faster than Pix4Dmatic for this alignment only aligned about 60% of the images (we used the Intel PC for the alignment of the images with Reality Capture). As can be seen in figure 3 below, most of the central area of the structure was not aligned. The settings used for this alignment were the following: max features per mpx - 10.000; max features per image - 40.000, Image overlap - medium; image downscale factor - 1; max feature reprojection error - 2; use camera priors for georeferencing - yes. Although the alignment process was not conclusive in this test, the capacity of Reality Capture to import the camera alignments from Agisoft Metashape (in Colmap format) allows to make use of the dense point cloud and mesh processing capacities of Reality Capture.

Figure 6. Alignment of the domes in Reality Capture

In Comparison, and as will be shown below Agisoft Metashape and Bentley Itwin Capture were able to align all images effortlessly at this stage. Taking into account the time constraints linked to the fact that this work was commissioned by UNESCO, we took the decision to move forward with the two photogrammetry applications that yielded the best results at this stage[1]. Therefore, we took the decision to focus solely on Agisoft Metashape and Bentley Itwin Capture for the complete alignment at this stage. The premise of the methodology for both software applications was the same: aligning first the outside of the structure; fixing its location with the use of GCPs; adding then room by room into the alignment; and finally building the mesh from all the aligned parts of the structure.

Bentley Itwin Capture

For the parametrization of the alignment process and modeling we based this work on Bentleys manual and recommendations[2] and followed the modeling process suggested by Bentley. We started with the "aerotriangulation"[3] of the outside area with the following settings : Adjustment constraints – none; final rigid registration – none; key point density – normal; pair selection mode – default; colour correction - machine learning; tie points – Compute; position – Extend; pre-calibration stage – Enabled. Once the initial alignment was done, we added the coordinates of the GCPs as "control points for adjustment". By doing so, we made sure that Itwin Capture would not realign the existing aerotriangulation and only "extend" it in the following process stage.

After this step we imported the remaining images of the inside of the structure (drone and DSLR images). In order to be able to align the images taken inside the domes which had poor GPS precision (sometimes several hundred meters off), we deleted the GPS data altogether from the metadata of the images and only kept the pitch/yaw/roll data of every image. We then proceed to extend the outside alignment with the images that were added.

Using this method, the software was able to align in these two steps 94% of all images (15441 out of 16371) in a process that took 14.5 hours. The 3D modelling stage per se at this point a straightforward process as it was after this stage a matter of creating the reconstruction making sure to choose adaptive tiling (as the ram usage would explode otherwise).

As the description of the methodology above shows, once the alignment process was done, the creation of the model was a matter of launching the process. As shown through the description of the process above, it was easy to achieve the 3D model with this software, and the average precision of the structure that was modelled is 0.6mm/px[4]. Figures 7 and 8 below provide a sense of the quality and sharpness of the result achieved with this software. Figure 7 (which depicts the passageway between circular rooms A and B) is of particular interest here as this image clearly shows the colour grading used by Bentley (which at times looks too saturated) and more importantly the sharpness of the bricks and plaster.

[1] This is not to say that PIX4Dmatic and Reality Capture could not be used to align the images and generate models with the acquired samples.
[2] The manual which includes recommendations on the best parameters is available online in Bentley's Resource Centre.
[3] Bentley Itwin Capture uses the term 'Aerotriangulation' to describe the initial alignment process of the images.
[4] This number comes from Bentley's quality report generated for the model

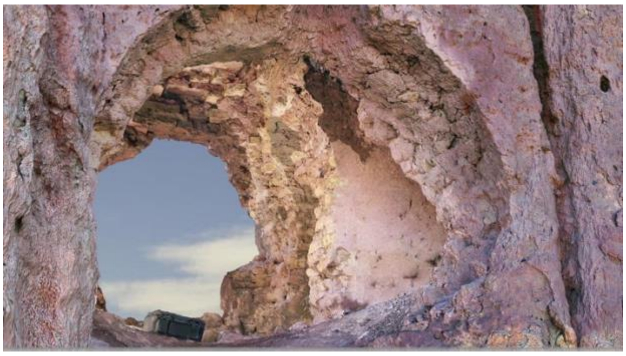

Figure 7. interior of a dome modelled with Bentley Itwin Capture

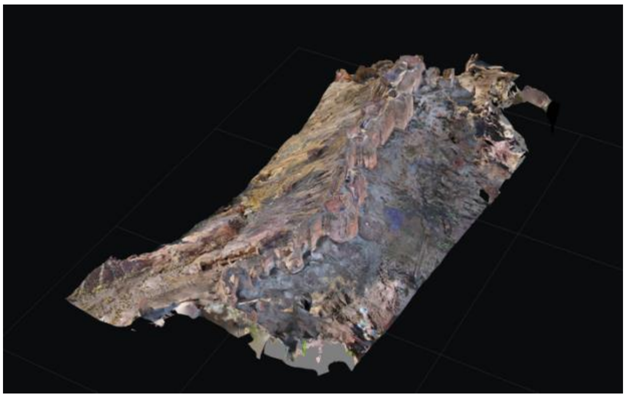

Figure 8. General view of the model generated with Bentley Itwin Capture

Taking into account all steps that were necessary to build the model with this software the list below provides a list of advantages and disadvantages encountered and noted throughout the process. The advantages of using Bentley Itwin Capture for this modeling exercice are the following:

- This software produces a sharp and high quality mesh/texture of the model (with an overall resolution of 0.6mm/px).
- The overall processing time was shortest with Itwin Capture (on our PC the alignment took 14.5h and the 3D reconstruction a little less than 4 days).
- The workflow is easy to master and does not include any complex steps after the aligment/aerotriangulation phase.
- Building a 3D model this large requires long processing times and computers can crash for diverse reasons during the rendering. Itwin Capture can continue at the point it was after a computer crash, thus reducing the need to restart a lenghty modelling process.
- Bentley Itwin Capture was able to align images taken with different cameras and focal lenghts from our DSLR and drones

The disadvantages that were encountered using Bentley Itwin Capture for this modeling exercice were the following:

- As it is not possible to edit splats/tie points as in Metashape or Reality Capture, the camera alignment cannot be manually edited or filtered. The only option in case of alignment errors is to realign the images with different parameters.-
- Bentley Itwin Capture is more expensive compared to Agisoft Metashape and requires a yearly subscription to get updates on the software.
- The colourgrading in this software uses machine learning and this leads often to overly saturated results[5].
- The software is only available on Windows
- Sharing the model not easy as viewing the Bentley native 3SM file requires a Bentley account (although this is free, it is not convenient to share a model to clients this way or publish it). Bentley Itwin Capture does allow OBJ exports, however, the software will cut the model in tiles and export each tile as a separate OBJ file that would need to be reassembled using external software like Blender.

Agisoft Metashape

As with Bentley Itwin Capture, the first step was to align the outside area first and to add the interior of the structure at a later stage. The alignment parameters in Metashape that worked best with the Skydio X10 and 2 were the following: accuracy – High; generic preselection – yes; reference preselection – source; key point limit - 60,000; tie point limit - 20,000; exclude stationary tie points – yes; guided image matching – no; adaptive camera model fitting – yes. With these settings Agisoft Metashape was able to align 100% of images taken with the drones[6]. Once the outside alignment was done, we cleaned the alignment using the parameters as shown below: Reconstruction Uncertainty – 20% of points; Projection Accuracy – 10% of points; Reprojection Error – 5% of points, making sure to optimize the camera alignment with all lens parameters checked and 'adaptive model fitting' checked after each step. Once this cleaning process was completed for the outside of the structure, we imported the GCP's with a view to 'anchor' this alignment and to avoid that by adding the images from the inside later, this alignment would be altered. Thus, we proceeded with the camera optimization of the outside alignment based on our 4 GCPs that were set in the vicinity of the site (making sure to uncheck 'adaptive camera model fitting' and turning on 'fit additional corrections').

For every room and passageway we created a separate chunk and made a separate alignment. And for each of these sparse point clouds we proceed with optimizing the camera alignment with the process described above. Using the align chunks function of Metashape (ensuring to have the 'fix scale' option turned on and making sure that the image matching quality is set to 'high' with a keypoint limit of 10.000) we proceeded to align the circular rooms and passageways into three separate chunks (one chunk for rooms and passageways A and B, one for

[5] We have contacted engineers from Bentley Itwin Capture on this issue and were told that this was an issue that they would fix.
[6] It should be noted however, that Agisoft Metashape was not able to align images added with our DSLR into the alignment.

https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-171-2025 | © Author(s) 2025. CC BY 4.0 License.

C and D, and one for F and G)[7]. After this step we aligned each of the 'inside' chunks separately with the alignment for the outside part of the structure. Only after all chunks were aligned did we proceed to merging the chunks[8]. The total processing time for the alignment of all chunks was 9 hours. Figure 9 below shows the sparse point cloud as seen in the software.

Figure 9. Tiepoints of the outside structure in Agisoft Metashape

At this stage the Alignment was cleaned again without however reoptimizing the cameras as this would have created problems with the merged chunks. The process as described above includes the parameters that have yielded most consistent results for the alignment and merging part. This was the most difficult aspect of the 3D modelling process.

For the following steps we followed Metashapes's recommendations as set out in the software manual. The process that was adopted consisted of creating the depth maps and dense cloud from the sparse point cloud, to filter the dense point cloud and to build a tiled model afterwards.

The depth maps and dense cloud parameters were set to high and the filtering mode to moderate (this took 3 days of processing time). It should be noted here that setting the overall quality to ultra-high would have resulted in all likelihood in a sharper texture[9]. Once this process completed, we filtered out points of the dense point cloud with a confidence level lower than 3. In the final stage of this modelling process we built the tiled model of the site based on depth maps on high quality (this took 1.5 days). The average precision of the structure that was modelled is 1.07mm/px[10].

In terms of pure processing time building the model with Agisoft Metashape took a little over 5 days. However, this does not account for setting parameters as well as filtering and cleaning the sparse and dense point cloud. Figure 10 below shows details of the passageway between circular rooms F and G and figure 11 the general view of the model. The sharpness and precision of the texture allows to see cracks and details on the bricks of the structure.

Figure 10. interior of a passageway between circular rooms F and G modelled with Agisoft Metashape

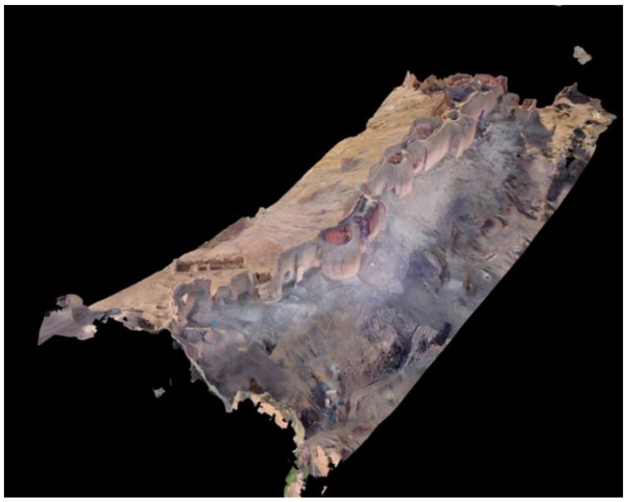

Figure 11. General view of the model generated with Agisoft Metashape

Just as with Bentley's Itwin Capture, this research has allowed us to better understand the advantages and disadvantages of using Agisoft Metashape for the construction of 3D models based on UAV imagery. The advantages of this software are the following:

- In this particular example, the color correction from Agisoft Metashape, albeit being less contrasted and sharp in comparison to the model from Bentley Itwin Capture is closer to the actual colours.

[7] We have created a separate chunk for room E using images of our DSLR. As Metashape could not automatically align and add these images we added this room manually by setting markers at the entrance.

[8] We found that it was not possible to redo the optimize camera alignment command at this stage, as doing so would consistently result in disrupting the alignment between the chunks.

[9] The machines we had at our disposal for this research did not have enough processing power to get through the processing of the depth maps and dense point cloud with these settings.

[10] This number comes from Agisoft Metashape's quality report generated for the model

https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-171-2025 | © Author(s) 2025. CC BY 4.0 License.

- Metashape gives the user great flexibility and access to filtering and editing the tie points of the initial alignment even manually. For a complex structure like Shahr-i Zohak this was very useful.
- Although Metashape is currently less popular in western governments and universities (due to the fact that this is a Russian Software), it remains commonly used and it easy to find solutions to problems online without having to contact the company.
- In terms of drone image alignment, Agisoft Metashape yielded the best results as the software was able to align 100% of the images taken with the two drones used. However it should be noted that the software was not able to add DSLR images to the alignment (although it was able to align the DSLR footage perfectly in a separate chunk).
- The native tiled model format (*.tls) can easily be shared and viewed either in the Agisoft viewer or on external 3D model viewing platforms like Nira. It also does not require an account which means that clients or the public can access the model via a simple link.
- Agisoft Metashape is available for Windows MacOS and Linux.

The disadvantages encountered by using Agisoft Metashape for this modeling exercice were the following:

- Processing times are lenghty and the creation of a model at the quality that we wanted to achive required many steps.
- Compared to Bentley Itwin Capture the resolution of the model is not as high (however it would be necessary to test the same model in Ultra High settings to compare the precision reached this way).
- When using Agisoft Metashape, a computer crash does not only mean that the current process is lost, it has also happened more than once that the whole directory was corrupted and that the software could not open the *.psx file anymore. It was therefore necessary to have backups of every stage of the work to prepare for this eventuallity.
- Agisoft Metashape requires a lot of harddisk space. In this case, the size of the directory that includes the sparse/dense point cloud, depth maps and tiled model is 2TB. This does not account for the extra space of the different chunks before alignment as well as the necessary backups to prevent loosing everything because of the computer of software crashing. In total the space required for this work in Metashape nearly reached 4TB.

Comparing the results against a benchmark

Both software applications that were used in this research have generated very high-resolution results (0.6 and 1.07 mm/px respectively). Figure 12 below presents a comparison of the quality of both software applications for the same area and adds as a benchmark actual footage from the Skydio 2 drone that scanned this particular circular room. As can be seen in this comparison, the model generated with Metashape better represents the violet and ochre tones of the lighting conditions, whereas the model generated with Itwin Captureis sharper and better shows the details of the different architectural features of the structure (bricks, mortar, decorative elements).

Benchmark Picture taken by the Skydio 2 drone of the passageway between circular rooms B and A

Screenshot of the same scene taken from the model achieved with Bentley Itwin Capture

Screenshot of the same scene taken from the model achieved with Agisoft Metashape

Figure 12. General view of the model generated with Agisoft Metashape

4. Conclusion

The results of this case study shows that using Skydio's X10 and S2 drones and setting a low Ground to surface distance (between 1 and 5 meters) and high vertical and horizontal overlap between the images (75%-95%) allowed us to achieve 3D models with an average accuracy of 1 millimetre per pixel (and even less using Bentley Itwin Capture) for a 120m long and 30m wide section of the fortified citadel of Shahr-i Zohak. In this comparative study, although both Metashape and Bentley generated excellent results, the easiness, simpler workflow, smaller file size and overall better result of Bentley's Itwin Capture would make working with this software much easier.

This research has also allowed us to clarify and test a methodology with clear workflows for both software applications, which is replicable, and which can be applied to other sites in similar conditions and contexts. It was equally our aim to be as transparent as possible with regard to the whole process and provide all 'paradata' as set out by the London Charter for the computer-based visualisation of Cultural Heritage.

Although this model cannot replace visiting the site, this model does allow for detailed and remote damage assessments and measurements of the archaeological site, and can be used in the context of Building information modelling application (BIM) by heritage professionals, conservation architects and archaeologists to understand the site and better focus their interventions (as the model is georeferenced and scaled with the help of Ground Control points that have been measured separately with the GNSS rover). However, it should be noted that the low GSD and high overlap (particularly for the inside areas) required more than 15.000 pictures to be taken for this area alone which is time consuming and process intensive at the alignment and modelling phase.

Acknowledgements

This research was conducted with the financial support of the Government of Japan through UNESCO. We are especially grateful to Prof. Kazuya Yamauchi from the University of Teikyo for his technical expertise and guidance. We also acknowledge the support of the engineer Amin Asifi who helped on site with the establishment of concrete benchmark points and verifying the accuracy of the measurements of the GNSS receiver with a total station. In addition to that we also acknowledge the support provided by Skydio regarding the optimal parameters for the X10 and S2 drones. Finally, we also acknowledge the support provided by Bentley Itwin Capture regarding the best parameters to use to merge different aerial and ground footage.

References

Baker, P.H.B., Allchin F.R., 1991: Shahr-I Zohak and the History of the Bamiyan Valley, Afghanistan. B.A.R. https://doi.org/10.30861/9780860547204

Margottini, C., Fidolini F., Iadanza, C., Trigila, A., Ubelmann, Y., 2015: The conservation of the Shahr-e-Zohak archaeological site (central Afghanistan): Geomorphological processes and ecosystem-based mitigation, Geomorphology, 239: 73–90. https://doi.org/10.1016/j.geomorph.2014.12.047

Robinson, M. G. P., 2024: Photogrammetry for Archaeological Objects: A Manual, Sydney University Press. https://doi.org/10.2307/jj.18108217

Küng, O., Strecha, C., Beyeler, A., Zufferey, J-C., Floreano, D., Dua, P., Gervaix, F., 2011: The Accuracy of Automatic Photogrammetric Techniques on Ultra-light UAV Imagery, UAV-g 2011 - Unmanned Aerial Vehicle in Geomatics. https://infoscience.epfl.ch/entities/publication/1f815a39-7a4b-4f4a-ad68-70b00a153f4f/conferencedetails

Bouledroua, D., 2025: 3D model of the domes A to H of the Triangular Plateau of Shahr-I Zohak. https://cloud.agisoft.com/shared/projects/da616ceb-e67c-43b9-a227-8f20fb6e6292

Denard, H., 2012: A New Introduction to the London Charter, in A. Bentkowska-Kafel, D. Baker & H. Denard (eds.) Paradata and Transparency in Virtual Heritage Digital Research in the Arts and Humanities Series. https://doi.org/10.4324/9781315599366

Agisoft Metashape User Manual: Professional Edition, Version 2.2, 2025, Agisoft LLC, St. Petersburg, Russia, Accessed: June, 20, 2025. [Online]. Available: https://www.agisoft.com/pdf/metashape-pro_2_2_en.pdf

Bentley Itwin Capture ressource center, 2025, Bentley Systems, Exton, Pennsylvania, USA, Accessed: June, 20, 2025. [Online]. Available: https://bentleysystems.service-now.com/community?id=kb_article_view&sysparm_article=KB0012289

Bentley Knowledge base, 2025, Bentley Systems, Exton, Pennsylvania, USA, Accessed: June, 20, 2025. [Online]. Available: https://bentleysystems.service-now.com/community?id=kb_view2&kb_id=4bcf16ac473102909091861f536d4351&spa=1&kb_knowledge_base=4bcf16ac473102909091861f536d4351

Pix4D Documentation, 2025, Pix4D, Lausanne, Switzerland, Accessed: June, 20, 2025. [Online]. Available: https://support.pix4d.com/hc/pix4dmatic#processing-options

RealityScan Help, 2025, Epic Games, Cary, North Carolina, USA, Accessed: June, 20, 2025. [Online]. Available: https://rshelp.capturingreality.com/en-US/startIndex.htm