

MULTI-SENSOR 3D RECORDING PIPELINE FOR THE DOCUMENTATION OF JAVANESE TEMPLES

A. Murtiyoso^{1*}, P. Grussenmeyer¹, D. Suwardhi², W.A. Fadilah², H.A. Permana² and D. Wicaksono²

¹ Photogrammetry and Geomatics Group, ICube Laboratory UMR 7357, INSA Strasbourg, France – (arnadi.murtiyoso, pierre.grussenmeyer)@insa-strasbourg.fr

² 3D Modelling and Information System, Remote Sensing and GIS Group, Bandung Institute of Technology, Indonesia – deni@gd.itb.ac.id, widiatmokoazis@gmail.com, (hafizh_ahsan, darmawaneka97)@students.itb.ac.id

Commission II, WG II/8

KEY WORDS: 3D Recording, Heritage, Multi-Sensor, Pipelines, Javanese temples

ABSTRACT:

3D recording is an important procedure in the conservation of heritage sites. This past decade, a myriad of 3D sensors has appeared in the market with different advantages and disadvantages. Most notably, the laser scanning and photogrammetry methods have become some of the most used techniques in 3D recording. The integration of these different sensors is an interesting topic, one which will be discussed in this paper. Integration is an activity to combine two or more data with different characteristics to produce a 3D model with the best results. The discussion in this study includes the process of acquisition, processing, and analysis of the geometric quality from the results of the 3D recording process; starting with the acquisition method, registration and georeferencing process, up to the integration of laser scanning and photogrammetry 3D point clouds. The final result of the integration of the two point clouds is the 3D point cloud model that has become a single entity. Some detailed parts of the object of interest draw both geometric and textural information from photogrammetry, while laser scanning provided a point cloud depicting the overall overview of the building. The object used as our case study is Sari Temple, located in Special Region of Yogyakarta, Indonesia.

1. INTRODUCTION

The recording of heritage buildings has seen much development in the last several decades. Classical drawings and maps are being replaced by 3D models, while traditional measuring techniques are enhanced by range-based and image-based sensors. However, this does not change the fact that heritage documentation remains a crucial effort in preserving them. This is more so with the ever-present threat of potential hazards, both anthropological (Fangi, 2019) and natural (Baiocchi et al., 2013). Nowadays, heritage documentation is often performed using photogrammetry, laser scanning, or the combination of both. Indeed, the multi-sensor approach is often employed in order to complement each technique's weaknesses (Murtiyoso et al., 2018). The study on the integration process of different sensors is also a very interesting topic which has been addressed by several researchers (Farella et al., 2019; Lachat et al., 2016).

Photogrammetry has shown to be a technique which is flexible and relatively easy to implement (Barsanti et al., 2014), while producing faithful 3D point clouds through the use of dense matching algorithms (Remondino et al., 2014). Textural data is also a highlight of photogrammetry, enabling the generation of photo-realistic 3D models. However, photogrammetry requires time to process correctly; even more so when hundreds and even thousands of images are involved (Murtiyoso et al., 2017b). In terms of data acquisition, it also requires a certain level of skill as to produce the best model possible, both in regards to geometric and textural quality.

Laser scanning data, on the other hand, provides a fast and generally accurate point cloud using its rapid ranging system. It

also generates a large amount of data, which sometimes becomes a problem of its own as the data redundancy is not necessarily of interest. This is more so in the case of complex objects such as heritage buildings and/or objects (Barsanti et al., 2014). Limitations on the textural quality of laser scanning are also observed in the literature, even when a camera is attached to the device (Hassani, 2015).

An integration of both techniques, in order to draw the respective advantages of each method and to complement their flaws, is therefore an interesting subject to address. Several approaches to data fusion was described in Magda Ramos and Remondino (2015), involving three levels of fusion, namely: (i) purpose-based, (ii) data-based, and (iii) dimension-based. This study performs a dimension-based data fusion which relies on georeferenced datasets. The common coordinate system of the data sources enables a direct superposition of the data into the same system, thus creating a hybrid 3D model where laser scanning data is used for the general view of the object while photogrammetry is used to represent the more detailed and texture-important parts (Murtiyoso et al., 2017a).

Within the framework of the Franco-Indonesian *Partenariat Hubert-Curien* (PHC) "Nusantara" and SAME-Perancis research project, a mission to document a few Javanese temples located in the provinces of Yogyakarta and Central Java was conducted in December 2018. Two monuments were recorded: a thorough and multi-sensor survey of Candi Sari (Figure 1) and an aerial photogrammetric survey of Borobudur temple complex. While Borobudur Temple represents a very unique example of Javanese temple architecture, for the purposes of this paper only Candi Sari will be used as case study due to the

* Corresponding author

immense scale of Borobudur. However, interested readers may also refer to a previous study regarding Borobudur Temple in Suwardhi et al. (2015).

During the mission, Candi Sari was recorded using a combination of laser scanning, terrestrial and UAV photogrammetry. A topographic survey (GNSS and traverse networks) was also conducted for this site in order to provide a 3D model georeferenced to the Indonesian national projection system. The topographic survey also measured check points in the interest of model accuracy assessment.

Javanese temples ("candi") are religious structures dating to the Middle Ages (9th to 15th century) and are either Buddhist or Hindu in nature. Most of these structures lie in the Indonesian provinces of Yogyakarta and Central Java, with some other notable examples existing in other provinces as well (Degroot, 2009). They are generally identified by the use of volcanic rocks as the main building material and frequently adorned by intricate bas-reliefs. Although these monuments are well documented in traditional mediums such as drawings, only few studies have been conducted on their 3D documentation from a geomatics point of view. Several examples of notable cases include the work of Hidayat and Cahyono (2016), Lehner (2017) and Suwardhi et al. (2015).

2. SITE AND INSTRUMENTS DESCRIPTION

The Javanese temple used as a case study in this paper is a 9th century Buddhist Candi Sari (Figure 1) located in Sleman Regency, Yogyakarta Special Region, Indonesia. The temple's dimensions are around 20 meters by 14 meters in a rectangular shape. The interior consists of three chambers with visible remains of a possible second story, giving the archaeological assessment that the temple was a *vihara* or a monastery for Buddhist monks rather than a pure place of worship (Degroot, 2009). The interior is also notable for its niches (Figure 2) that were possibly intended for images of *garbha-dhatu* deities, as is the case of Candi Mendut, another Javanese-style Buddhist temple¹. Some sculptures adorn the exterior façade, including an imposing sculpture of the traditional *kala* above the entrance and the niches in the interior.

The laser scanning data was obtained using a terrestrial laser scanner (TLS) Faro Focus M70. This TLS is a phase-based scanner designed for a close range acquisition (from 0.6 m up to 70 m as declared by the manufacturer). It is therefore more suitable for interior and close range applications. It has a theoretical precision of 2 mm for an object located at a distance of 10 m². In order to be able to generate 3D models of nearer objects, photogrammetry was conducted using both a DSLR camera and a drone to retrieve images of the temple roofs as well as oblique images. The DSLR used was a Canon 5D Mark IV (30 MP), while the DJI Phantom 4 was used to acquire the drone images. In addition to this, topographical survey was also conducted using a total station and GNSS receiver. This was done by establishing a traverse network around the temple, attached to the Indonesian national projection system by 2 GNSS reference points. Both the TLS and photogrammetry data were then georeferenced using this topographical information in order to fuse them in the same system.



Figure 1. Candi Sari as photographed from its exterior.



Figure 2. One of the niches in the interior of the temple.

3. ACQUISITION PIPELINE

A preliminary survey was conducted prior to the mission. This involves a rough sketch of the temple surroundings, in order to determine points for the traverse network as well as fly paths for the drone. This was conducted using Google Earth tool (Figure 3). In this case, the site is located in an open field and therefore poses little problem in terms of vegetation cover and masking for GNSS and drone measurements. The only problem was the legal matter regarding drone flight on and around the temple, as it is normally forbidden. A special permit had to be issued in order to be able to use the drone for the data acquisition.

The subsequent pipeline employed in the surveys follows a multi-sensor approach (Murtiyoso et al., 2018) which aims to integrate the results of the various acquisition methods. In the case of Candi Sari, laser scanning was performed to obtain the interior of the temple while serving as a "bridge" between the interior and the exterior which was acquired using photogrammetry (both terrestrial and UAV). Several carvings inside the temple were also documented using photogrammetry, in order to render them in a higher resolution. The laser scanning data was processed using FARO Scene, while photogrammetric work was done using the software Agisoft Metashape. A traverse network was established around the temple; with reference points measured using static GNSS. The topographical network was measured in the national cartographic system. The integration of the sensors was performed by using the control points scattered around and inside the temple in the form of coded targets and artificial spheres. The artificial spheres were standardised spheres issued by Faro, with a uniform diameter of 13.9 cm.

¹ <https://www.britannica.com/art/Southeast-Asian-arts/> retrieved on 20 June 2019

² <https://insights.faro.com/long-range-laser-scanners/techsheet-faro-focus-s-m-laser-scanner> retrieved on 20 June 2019

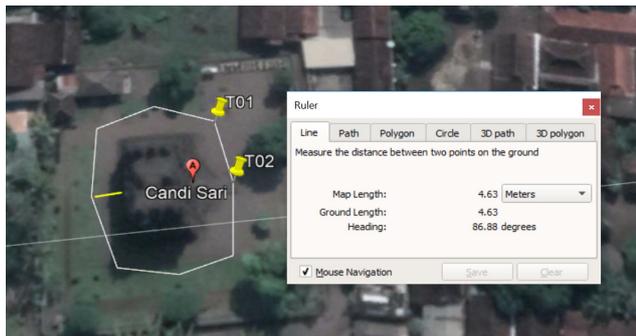


Figure 3. Preliminary planning using GoogleEarth.

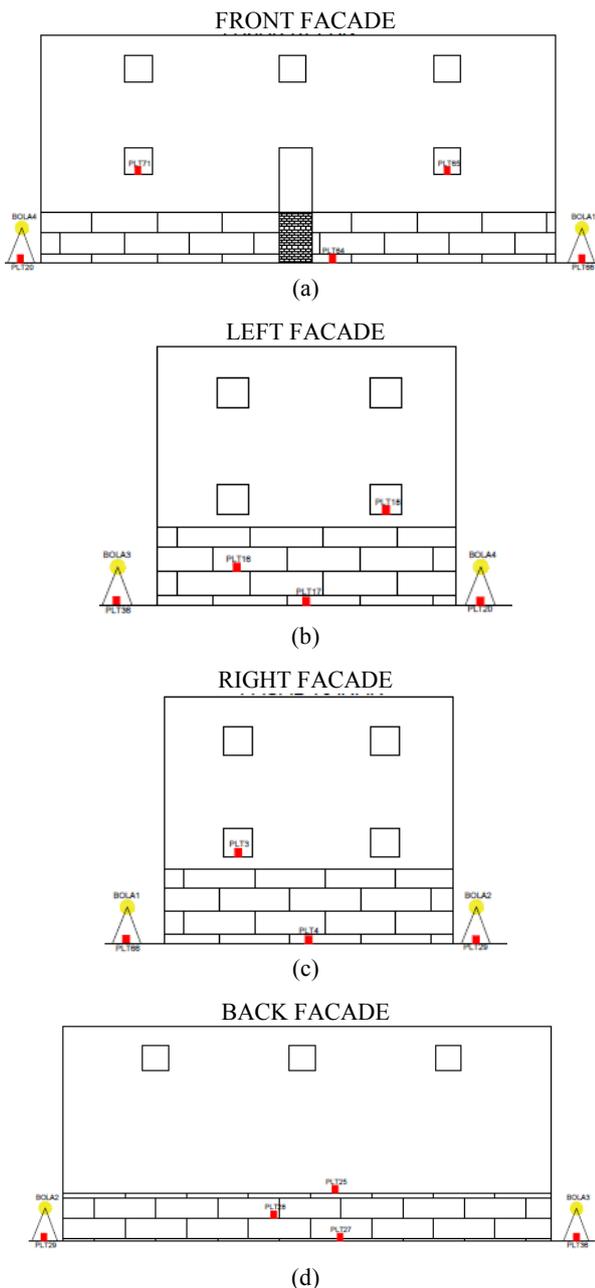


Figure 4. Sketch of the temple's façades: (a) front, (b) left, (c) right, and (d) back. The red rectangles denote detail points (PLT#) measured using the total station, while the yellow spheres denote the artificial spheres (BOLA#) used in TLS registration, also known in coordinates.

These supports were thereafter measured using the total stations from at least two stations, thus assuring a spatial intersection for each target. In addition about a dozen detail points were also measured on the front façade of the temple, to serve as check points in the eventual accuracy assessment. The topographical points served as a common reference for both the photogrammetry and laser scanning results, integrating them into the same system automatically after their respective georeferencing steps.

Javanese temples present particular challenges that require adaptations to the established pipeline. The absence of any lighting in the temple interiors necessitate the use of intensity images instead of RGB images for the laser scans. The complex nature of the reliefs also meant that in many cases terrestrial laser scanning was not enough to represent them. We resorted to close range photogrammetry to obtain the required resolution for these cases. The tropical climate surrounding both sites also meant that particular care should be taken for the tools used, while the presence of tourists often creates masks in the resulting point cloud, if not delays during the acquisition phase.

4. RESULTS AND DISCUSSIONS

4.1 TLS Processing

The laser scanning technique generates point clouds in the relative coordinate systems of each of the acquiring station. The TLS processing involves therefore the fusion of these different scan stations into a single entity in the same coordinate system. The Faro TLS used in this study employs an indirect georeferencing approach, whereas a prior registration process was performed to integrate all the scan stations in a single, albeit relative, space. This step is then subsequently followed by the georeferencing step by introducing the real-world coordinates of some points in the point cloud. The TLS processing was performed using the software Faro Scene which is developed by the TLS manufacturer.

The registration was performed using the artificial spheres as common points. As can be seen in Figure 4, one sphere was placed on each of the four corners of the temple exterior. This is in order that each scan station contains at least three spheres as common points. Whenever the spheres did not provide good results, the detail points were used as auxiliary points to aid the registration. A total of 8 scan stations were taken for the temple exterior; 4 in front of each façade, and another 4 at each corners. The registration yielded a result of 4.3 mm in maximum residual error and 3.8 mm in average residual error between the common points. In terms of overlap, the minimum overlap between the scan stations was 10.20%.

The artificial spheres were also measured using total station, and therefore known in coordinates. The georeferencing was performed using the spheres, because the spheres were able to be detected automatically by Faro Scene and therefore presented better precision compared to the detail points. In this regard, the spheres were distributed evenly around the object of interest, in order to ensure a homogeneous result. The spheres' coordinates were measured and computed by applying an offset of 7 cm which corresponds to its radius.

In total, four sphere coordinates were used in the georeferencing (Table 1). The RMSE error amounts to 5.7 cm, which indicates the presence of some systematic error. This may be either due to the quality of the sphere coordinates, or the quality of the detection of the sphere centres.

Name	Easting (m)	Northing (m)	Alt. (m)	Error (m)
BOLA1	442032.699	9142041.91	164.721	0.060
BOLA2	442018.342	9142042.26	164.620	0.071
BOLA3	442018.065	9142021.93	164.632	0.053
BOLA4	442031.914	9142022.01	164.713	0.041
		RMSE		0.057

Table 1. Residual table for the georeferencing step of the TLS data using the four artificial spheres.

Name	Easting (m)	Northing (m)	Alt. (m)	Error (m)
PLT4	442025.124	9142042.159	163.807	0.086
PLT16	442022.540	9142022.815	165.320	0.008
PLT18	442027.512	9142023.482	167.462	0.023
PLT27	442018.187	9142033.284	163.670	0.018
PLT28	442019.799	9142036.310	165.133	0.037
PLT29	442018.370	9142042.201	163.692	0.044
PLT64	442033.039	9142036.611	163.700	0.031
		RMSE		0.042

Table 2. Residual table for the check points for the registered and georeferenced TLS data.

In order to assess the accuracy of the registered and georeferenced TLS data, a comparison of coordinates between those of the point cloud and those of the detail points as measured by total station was conducted. A summary of this analysis can be seen in Table 2. The order of the RMSE values is more or less similar to that of the RMSE of Table 1, specifically 4.2 cm.

4.2 Photogrammetric Processing

There are two data sources for the photogrammetry part of this study. The first one involves aerial images as taken by a drone. The main objective of the drone images was to complete data of the object from otherwise inaccessible parts such as the roofs. Another objective of the drone data was to complement the terrestrial photogrammetry data by taking oblique images. The second data source is therefore terrestrial photogrammetry using a DSLR camera. Both data sources were processed using Agisoft Metashape in two different "chunks", which were then fused using a similar concept as the one used later for integration with TLS data (georeferencing to a common system). Over 500 drone images were taken, while the amount of DSLR images was also around 500 images.

The drone images were taken from variable flying heights of 5 m, 10 m, and 20 m to create a "dome" of concentric rings. The terrestrial images were on the contrary taken from a distance of 3-5 m from the object. In terms of bundle adjustment results, the drone data yielded an RMSE value of control points of 1.9 cm as computed from 4 control points. Terrestrial processing used more control points (6), giving an RMSE value of 3.3 cm. Check points were also marked on both data sources. In this regard, a total of 17 points were used as check points for the drone data, while 7 points were used for terrestrial data, with RMSE values of 3.8 cm and 1.3 cm respectively.

The results of the bundle adjustment were deemed sufficient and dense matching was subsequently applied to obtain the dense point cloud from both drone and terrestrial data. A similar analysis as the one showcased at Table 2 was performed on the combined dense point cloud of the photogrammetric data, as can be seen in Table 3. The overall RMSE of the check point residuals gave a value of 5 cm, which once again concurs with the order of value of the previous numerical analyses. The photogrammetric point cloud was then integrated with the TLS point cloud in order to generate a complete and hybrid model.

Name	Easting (m)	Northing (m)	Alt. (m)	Error (m)
PLT4	442025.124	9142042.159	163.807	0.080
PLT16	442022.540	9142022.815	165.320	0.007
PLT18	442027.512	9142023.482	167.462	0.031
PLT27	442018.187	9142033.284	163.670	0.029
PLT28	442019.799	9142036.310	165.133	0.050
PLT29	442018.370	9142042.201	163.692	0.057
PLT64	442033.039	9142036.611	163.700	0.030
		RMSE		0.050

Table 3. Residual table for the check points for the photogrammetric point cloud.

4.3 Data Integration

The integration of the TLS and photogrammetric data was aimed to take the best part of each data set and to create a complementary final result. As a consequence, prior to the integration itself, segmentation on both data (TLS and photogrammetry) was performed for some area of interest such as statues and carvings (Figure 5).

From both samples of segmented part, a density analysis was performed using the CloudCompare software. In this analysis, a sphere with a radius of 10 cm was created for each point within the point cloud, and the number of points contained within this sphere was then annotated to each point to give an illustration to the density of the data. As can be seen in Figure 6, the TLS data showed a severe lack of density in many parts (red zone) compared to the same analysis applied to the photogrammetric data. From this analysis, the initial hypothesis that photogrammetric dense cloud is better suited for detailed parts was validated.



Figure 5. Example of a carving on the temple façade, resulting from (a) photogrammetry and (b) TLS.

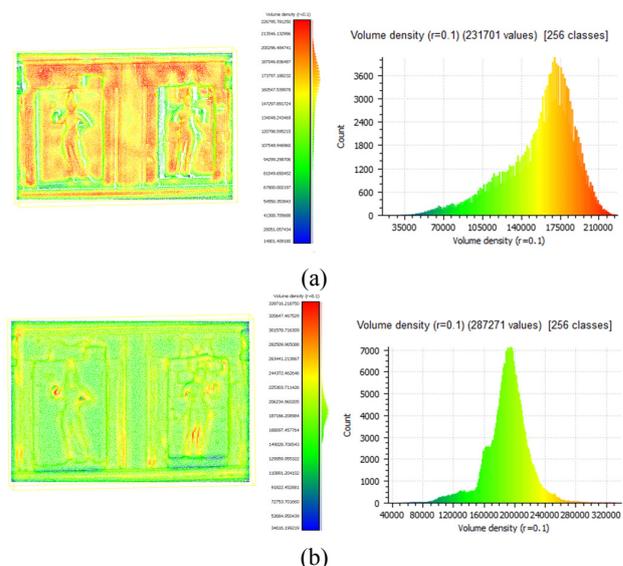


Figure 6. Density analysis for the carvings of Figure 5: (a) density of TLS data and (b) density of photogrammetry data.

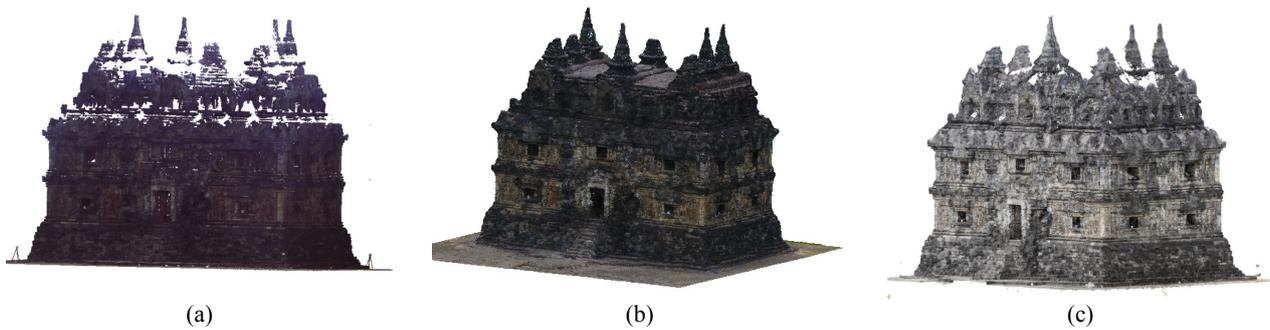


Figure 7. Dense point cloud from each of the data source: (a) TLS, (b) drone, and (c) terrestrial photogrammetry.

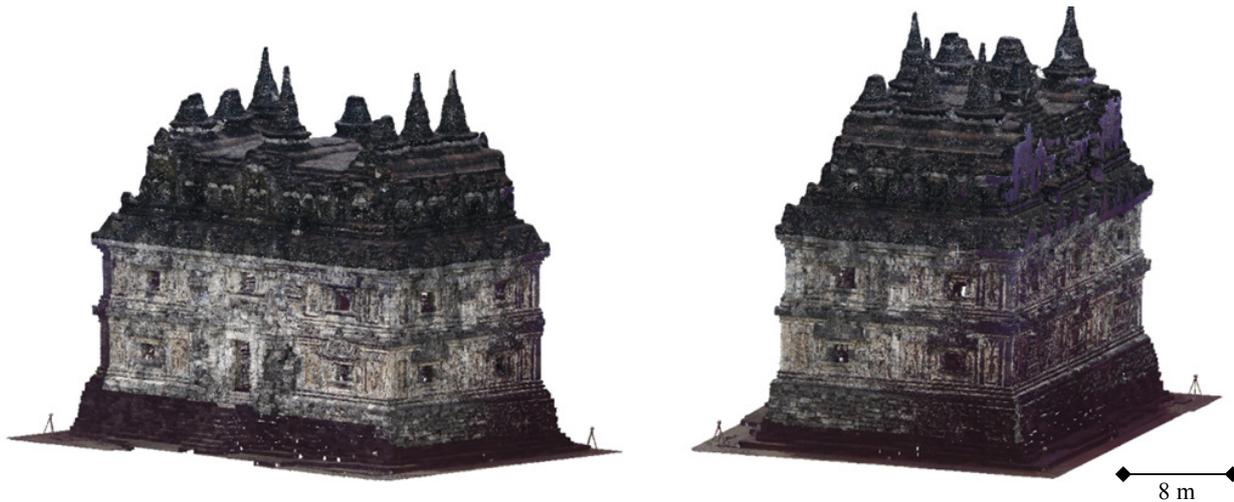


Figure 8. Illustration of the resulting fused point cloud.

As can also be seen in Figure 7, each data source possesses its own advantages and is therefore complementary as expected. While terrestrial photogrammetry provides the best density in terms of point cloud due to its closer sensor-to-object distance, the processing of over five hundred 30 MP images takes a very long time. Meanwhile, not all parts of the temple require this level of density. Namely, the statues and carvings may benefit from a denser point cloud, but simple walls may find that TLS suffices. This knowledge is important since it may save time during the acquisition step. The drone data, meanwhile, remains lower in terms of density even compared to TLS data due to its sensor specifications. However, drones are vital in filling the holes left by terrestrial techniques, both TLS and photogrammetry as evidenced in Figure 7.

The final integrated point cloud is illustrated in Figure 8. The segmentation step enabled to avoid double points, therefore reducing the final file size. In terms of fusion, the three data sets were successfully fused using the mutual georeferencing approach.

5. CONCLUSION AND FUTURE WORK

Several products were generated from the resulting point cloud, including textured 3D mesh models and orthophotos. Another interesting result was the virtual reality environment for the laser scanning point cloud generated using the Faro Scene software. This virtual reality was created from the registered TLS laser scans and serves as a very interesting medium to disseminate information regarding these historical sites. Other deliverables include a 3D print of some of the sculptures and reliefs of the temples.

The 3D documentation pipeline implemented in this project involves the integration of data resulting from various sensors. Care was also taken to ensure precision of the results, all while considering the different challenges and conditions of documenting Javanese temples. While the terrestrial laser scanner (TLS) is very valuable in producing 3D point cloud, stations were required to be placed close enough to the object in order to generate high precision results. In this regard, photogrammetry provides a better solution. However, the TLS is still useful in providing overall 3D models and “bridges” between the different photogrammetric projects. Integration of these two techniques is therefore a very valuable solution to thoroughly map monuments which such high levels of detail. This is the case for intricate Javanese temples, but a similar reasoning can also be applied to other Southeast Asian architecture or indeed any detail-rich structures.

More studies are still planned for this research project. At the current level, the processing and analysis has not included the interior of the temple yet. The next improvement would be to integrate point clouds of the interior and the exterior using the same approach, since several control points were also measured in the interior using total station. The interior of Candi Sari is also interesting due to its ornately carved niches; an example of the dense point cloud of such object can be seen in Figure 9.

Another research which is conducted in parallel involves the use of the PolyFit algorithm (Nan, 2017) to try to generate polygonal shapes automatically from the Candi Sari data set. This is done with the objective of facilitating the creation of a 3D GIS system. A preliminary result of a simple polygonal shape is shown in Figure 10.



Figure 9. Example of the dense point cloud of the ornate niche in the temple interior

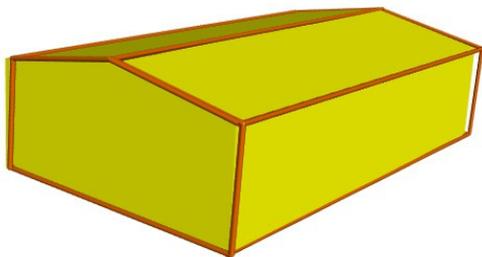


Figure 10. Preliminary result of the application of the PolyFit algorithm on the Candi Sari data set.

ACKNOWLEDGEMENTS

This research is part of the Franco-Indonesian Partenariat Hubert-Curien (PHC) NUSANTARA program under the aegis of the Indonesian Ministry of Research and Higher Education (KEMENRISTEKDIKTI), the French Ministry for Europe and Foreign Affairs (MEAE), and the French Ministry of Higher Education, Research, and Innovation (MESRI). The authors also wish to thank our colleagues in the Yogyakarta Archaeological Bureau, namely Sugeng Riyanto, and the Central Java Heritage Conservation Bureau for their help in facilitating the acquisition mission and in obtaining the necessary authorisations.

REFERENCES

Baiocchi, V., Dominici, D., Mormile, M., 2013. UAV application in post-seismic environment. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XL-1/W2, pp. 21–25.

Barsanti, S.G., Remondino, F., Fernández-Palacios, B.J., Visintini, D., 2014. Critical factors and guidelines for 3D surveying and modelling in Cultural Heritage. *International Journal of Heritage in the Digital Era* 3, 141–158.

Degroot, V., 2009. *Candi, Space and Landscape: A Study on the Distribution, Orientation and Spatial Organization of Central Javanese Temple Remains*, Mededeling. ed. Sidestone Press, Leiden.

Fangi, G., 2019. Aleppo - Before and after. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W9, pp. 333–338.

Farella, E.M., Torresani, A., Remondino, F., 2019. Quality Features for the Integration of Terrestrial and UAV Images. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W9, pp. 339–346.

Hassani, F., 2015. Documentation of cultural heritage techniques, potentials and constraints. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XL-5/W7, pp. 207–214.

Hidayat, H., Cahyono, A.B., 2016. Digital Reconstruction of Singosari Temple Using Structure From Motion Methods. *Geoid* 11, 211–218.

Lachat, E., Landes, T., Grussenmeyer, P., 2016. Combination of TLS point clouds and 3D data from kinect V2 sensor to complete indoor models. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Vol. XLI-B5, pp. 659–666.

Lehner, E., 2017. Advanced survey of ancient Buddhist and Hindu temples in central Java for tracking their position within the history of Southeast Asian architecture. *Journal of Comparative Cultural Studies in Architecture* 10, 21–28.

Magda Ramos, M., Remondino, F., 2015. Data fusion in cultural heritage - A review. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XL-5/W7, pp. 359–363.

Murtiyoso, A., Grussenmeyer, P., Guillemain, S., Priloux, G., 2017a. Centenary of the Battle of Vimy (France, 1917): Preserving the Memory of the Great War through 3D recording of the Maison Blanche souterraine. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. IV-2/W2, pp. 171–177.

Murtiyoso, A., Grussenmeyer, P., Suwardhi, D., Awalludin, R., 2018. Multi-Scale and Multi-Sensor 3D Documentation of Heritage Complexes in Urban Areas. *ISPRS International Journal of Geo-Information* 7, 483.

Murtiyoso, A., Koehl, M., Grussenmeyer, P., Freville, T., 2017b. Acquisition and Processing Protocols for UAV Images: 3D Modeling of Historical Buildings using Photogrammetry. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. IV-2/W2, pp. 163–170.

Nan, L. (2017). *PolyFit: Polygonal Surface Reconstruction from Point Clouds*. Visual Computing Center, KAUST.

Remondino, F., Spera, M.G., Nocerino, E., Menna, F., Nex, F., 2014. State of the art in high density image matching. *The Photogrammetric Record* 29, 144–166.

Suwardhi, D., Menna, F., Remondino, F., Hanke, K., Akmalia, R., 2015. Digital 3D Borobudur - Integration of 3D Surveying and Modeling Techniques. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XL-5/W7, pp. 417–423.