Identification of photo number effect for 3D modeling in Agisoft software

Monica Maharani$^{1*}$, Anna Charieninna$^1$, Hary Nugroho$^1$

$^1$Department of Geodesy Engineering, National Institute of Technology 23 PH.H. Mustofa Street, 40124 Bandung, Indonesia

e-mail: mahrani.m@itenas.ac.id

Abstract. At present, photogrammetry projects, including Close-Range Photogrammetry, has been supporting by software that works automatically. Agisoft Metashape is one of the Close-Range Photogrammetry data processing software that can work automatically. In accordance with ISPRS recommendations, the success of Close-Range Photogrammetry processing is determined by the overlap in each photo. Based on ISPRS, Close-Range Photogrammetry will succeed if each photo has an overlap of 90-95%. In line with ISPRS, Agisoft always recommends users to input more than enough photos in each processing, where more photos will produce better results. These conditions will make the number of photos used in Close-Range Photogrammetry processing will be very large, very long time processing and the processors required must be highly certified. This paper will analyze the photo number effect for 3D modeling in Agisoft software, as well as prove whether large overlapping values and a large number of photos will increase the accuracy of the resulted model. There are three simulations of photo number will be compared in terms of geometric quality, while the suitability of the resulting 3D model will also be discussed. Furthermore, the detection of causes failure in one simulation of photo number will also be explained in this paper. Results show that a large number of photos did not produce a better 3D model, so to create a 3D model does not require a large number of photos but the optimal number.

1. Introduction

One method that can be used to document 3D models is Close-Range Photogrammetry [1, 2]. Close-Range Photogrammetry represents photogrammetry that uses photos as a measuring instrument with a distance of less than 300 meters [3, 4]. In line with the times, 3D modeling using Close-Range Photogrammetry is supported by the existence of data processing software that works automatically. One of the software that runs automatic close-range photogrammetry data processing is Agisoft Metashape. Agisoft Metashape is an advanced image-based 3D modeling solution aimed at creating professional quality 3D content from still images [5]. Agisoft Metashape works with the latest multi-view 3D reconstruction technology, which is structure from motion. It works by imported photos taken from any position, as long as the object to be reconstructed is seen in at least two photos. To provide a guarantee of good processing results, Agisoft advised making the number of photos more than required because it's better than not enough. Moreover, based on ISPRS [6] one of the key stages which allows for accurate spatial measurement using digital cameras for Close-Range Photogrammetry is obtain images for object recording using a convergent image pair configuration with an overlap of 90-95%. Therefore, to make the number of photos more than needed, Agisoft's processing software users always make a lot of photos with large overlap. However, with large overlap, the number of photos produced is large so it is not effective in processing, because it makes a long processing time and large processing data. Long processing time and the large data size certainly make a decrease in work effectiveness, especially in production. This paper will analyze the photo number effect for 3D modeling in Agisoft software, as well as prove whether large overlapping values and a large number of photos will increase the accuracy of the resulting model. Several simulations of the

*Corresponding author.
photo number used in processing. The highest photo number that covers all objects with 90% overlap is reduced while keeping in mind the overlapping values of each photo is same.

2. Methodology
A case study was performed on Bosscha Observatory in Lembang, Bandung Barat Districts, West Java, Indonesia, which was recently inscribed in the Indonesian Cultural Heritage list. Bosscha Observatory can be seen in Figure 1. Bosscha Observatory has dynamic dome-shaped roofs and cylindrical building walls. The Bosscha Observatory wall is white with no motif, so it is almost homogenized on all sides.

![Figure 1. Exterior model of Bosscha Observatory generated by UAV 3D processing data.](image)

2.1. Data processing
The formation of a 3D model uses two simulations namely the first simulation using as many photos as possible and the second simulation uses a small number of photos while still paying attention to one photo with another overlapping. Using Agisoft, a 3D model can be generated in a fully automated three-step process [5, 7], comprising (i) the alignment of the photographs, (ii) the calculation of a dense 3D surface and finally (iii) the texture mapping of the model. The processing of the 3D model starts by aligning the photographs. During this step (i) a 3D sparse point cloud is generated representing the geometry of the scene. For the processing of the alignment, Agisoft uses a structure from motion approach [8, 9], (ii) The next step is absolute orientation or georeferencing by input the ground control points coordinate. And then doing for (iii) the relative orientation of camera position at the moment of image acquisition is determined and (iv) the internal camera parameters (focal length, principal point location, skew, radial and tangential distortion coefficients) are computed. (v) Generating dense point cloud, to produce a very tight point. After a dense point cloud is generated, the next step is (vi) making mesh, which is to collect 3D from point cloud resulting from the process of building dense point clouds. The concept of mesh formation is the result of the capture of a triangular net from dots formed from dense cloud. This net creates interrelated polygons creating a surface from a 3D model. This stage is done to bind a collection of cloud that has not been arranged, so that they close each other. (vii) The last step is making 3D models in this research is texture making. This stage aims to give texture and colour to the 3D model so that it can approach the actual state of the object. Parameters chosen based on recommended settings by software developers and test-and-trial (Table 1.)
| Photo Alignment Parameters          |            |
|-------------------------------------|------------|
| Accuracy                            | High       |
| Reference preselection              | Enabled    |
| Point Limit                         | 274.728    |
| Camera Accuracy                     |            |
| Marker Accuracy                     |            |
| Scale Bar Accuracy                  |            |
| Projection Accuracy                 |            |
| Tie Point Accuracy                  |            |
| Fit $f$                             | Enabled    |
| Fit $cx$, $cy$                      | Enabled    |
| Fit $k1$                            | Enabled    |
| Fit $k2$                            | Enabled    |
| Fit $k3$                            | Enabled    |
| Fit $k4$                            | Enabled    |
| Fit $b1$                            | Enabled    |
| Fit $b2$                            | Enabled    |
| Fit $p1$                            | Enabled    |
| Fit $p3$                            | Enabled    |
| Fit $p4$                            | Enabled    |
| Building dense point cloud          |            |
| Quality                             | High       |
| Depth Filtering                     | Mild       |
| Calculate point colours             | Enabled    |
| Building mesh                       |            |
| Source data                         | Dense Cloud|
| Surface type                        | Arbitrary (3D) |
| Face count                          | High       |
| Calculate vertex colours            | Enabled    |
| Building Texture                    |            |
| Mapping Mode                        | Generic    |
| Blending Mode                       | Mosaic     |
| Texture size/ count                 | 8192 x 1   |
| Enable hole filling                 | Enabled    |
| Enable ghosting filter              | Enabled    |
3. Results
A total of 362 dan 270 overlapping photos were used to reconstruct the 3D model (Figure 2 (left) dan Figure 2 (right)). Models that use 362 photos require almost 28 hours for scratching. Meanwhile, 270 photos take 13 hours. Regarding the accuracy produced, the root mean square error for the 3D model, as calculated from the subset of 38 ground control points, can be seen in Table 2.

Table 2. Comparison Of RMSE Values On Both Models

| Model                      | X error (m) | Y error (m) | Z error (m) | Error (pix) |
|----------------------------|-------------|-------------|-------------|-------------|
| Model with 362 photos      | 0.009       | 0.007       | 0.006       | 3.382       |
| Model with 270 photos      | 0.009       | 0.006       | 0.005       | 1.342       |
| Deviation                  | 0           | 0.001       | 0.001       | 2.04        |

Ground Sample Distance (GSD) obtained by both models can be seen in Table 3.

Table 3. Comparison Of GSD Values On Both Models

| Model                      | GSD (mm/pix) |
|----------------------------|--------------|
| Model with 362 photos      | 1.67         |
| Model with 270 photos      | 1.7          |
| Deviation                  | 0.03         |

Figure 2. Model with 362 photos (left) and Model with 270 photos (right)
4. Discussions

The 3D model produced from 362 photos with a total of 18-20 photos per flight line can be seen in Figure 3 (left). To make a 3D model with 270 photos several experiments were carried out, the first experiment was by reducing half the photo portion of the previous 360 photos 3D model which was 8-10 photos per line. The consideration of choosing photos to be 8-10 photos per flight line is the removal of photos in the overlapping part. Photos are reduced to too much overlap while keeping in mind that all photos on a flight line must include Bosscha buildings from the roof to the ground. Then, processing data with a number of 8-10 photos per line is carried out, but matching fails occur. This is because there are parts of the building that do not overlap between photos. Therefore, one and two more photos were added in one flight line to cover the hollow part but still failed to match. This matching failure can be seen in Figure 4. After the experiment, the most optimal 3D model covering all parts of Bosscha building is to use photos from 10-13 photos per flight line and the number of photos with 270 overlaps that vary from the flight line (Figure 3).

![Figure 3. Matching Failure](image)

A 3D model made of 270 photos (3D B Model) that processes it using the same parameters each time it produces GSD of 1.7 mm/pix. This value is not much different from the 3D model built from 362 photos (3D A model) which has a GSD value of 1.67 mm/pix (Table 2.). The GSD value that has a minimal difference is caused by the camera and photos used to replace the two 3D models are the same. Visually, the results of the 3D A model have a more similar appearance to the 3D B model and do not have as much noise as the 3D model A. However, in some parts there is a difference in the right side of the pipe. The 3D A model has a straight shape while the 3D B model has a slightly dented shape. Even though the 3D detail model presents more real conditions, the model B visual clarity is better. The comparison of the two models can be seen in the Figure 2.

GCP points used for the georeferencing process in the 3D B model are the same as the GCP points used in the 3D A model, which are 38 pieces. If seen from the geometry accuracy value generated from the processing stages of the 3D B model, the value of the Root Mean Square Error (RMSE) based on the GCP point entered and projected on the photo in the 3D model B gets a value smaller than the 3D model A (Table 1.) Because when the GCP marking photo in model A can be at the point of marking it is a point of noise. This noise point is based on the projection results with levelling from some of the actual shooting results that are wrong.
5. Conclusions
To make a 3D model, the photos needed do not need much, but the number of photos is optimal so that each photo must match those that overlap each other.

6. Acknowledgments
The authors are grateful for funding from PT. Zenit Era Utama Servizio (ZEUS).

7. References
[1] Yilmaz, H.M., et al., Importance of digital close-range photogrammetry in documentation of cultural heritage. Journal of Cultural Heritage, 2007. 8(4): p. 428-433.
[2] Park, H.S., et al., A new approach for health monitoring of structures: terrestrial laser scanning. Computer- Aided Civil and Infrastructure Engineering, 2007. 22(1): p. 19-30.
[3] GJ, I.N.H., PEMBUATAN MODEL TIGA DIMENSI CANDI GEBANG MENGGUNAKAN METODE FOTOGAMETRI JARAK DEKAT. 2014, Universitas Gadjah Mada.
[4] Luhmann, T., Robson, S., Keyle, S., Harley, Ian, Close Range Photogrammetry Principles, Techniques dan Applications. 2006.
[5] AgiSoft, L., AgiSoft PhotoScan User Manual: Standard Edition, Version 0.8. 4. AgiSoft LLC, 2011.
[6] V, I.-C., Tips For The Effective Use Of Close Range Digital Photogrammetry For The Earth Sciences. 2010: p. Close-Range Sensing: Analysis and Applications.
[7] Verhoeven, G., Taking computer vision aloft–archaeological three-dimensional reconstructions from aerial photographs with photoscan. Archaeological prospection, 2011. 18(1): p. 67-73.
[8] Ullman, S., "Interpretation of structure from motion"Proceedings of the Royal Society of London. Series B, 1979. 203: p. 405-426.
[9] Szeliski, R., Structure from motion, in Computer Vision. 2011, Springer. p. 303-334.