Comparative Analysis of the Semantic Conditions of LoD3 3D Building Model Based on Aerial Photography and Terrestrial Photogrammetry

Muh.Apriansyah[1,*], Harintaka[2]

[1] Departemen of Civil Engineering, Faculty of Engineering, Universitas Muhammadiyah Bima, Jl.Anggrek No 16 Ranggo Na'e Rasanae Barat Kota Bima, Indonesia.
[2] Departemen of Geodesy, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia.
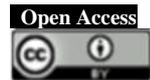*Corresponding author email: muh.apriansyah01@gmail.com

Abstract

3D modeling of buildings is an important method in mapping and modeling the built environment. In this study, we analyzed the differences between the semantic state of actual buildings and 3D models of LoD3 buildings generated using aerial and terrestrial photogrammetric methods. We also evaluated the accuracy of the visual representation as well as the suitability of the building geometry and texture. Our method involves collecting aerial and terrestrial photographic data and processing it using SFM (structure from motion) technology. The photogrammetric data was then processed using image matching algorithms and 3D reconstruction techniques to generate 3D models of LoD3 buildings. The actual semantic state of the building was identified through field surveys and reference data collection. The 3D building model was successfully modeled from 1201 photos and 19 ground control points. The results of the evaluation of the geometry accuracy test, dimensions and semantic completeness of the 3D model, the use of aerial photographs and terrestrial photogrammetry in LoD3 3D modeling are assessed from the results of the automatic 3D modeling process using SfM (Structure from Motion) technology that produces 3D building models in Level of Detail (LoD) 3 with Root Mean Square Error values <0.5 meters and has semantic completeness of the building in accordance with the original object based on the City Geography Markup Language (CityGML) standard. The facade formed from the modeling almost follows the original model such as doors, windows, hallways, etc.

Keywords: 3D, LoD, Semantic

1. Introduction

In recent decades, the development of mapping and photogrammetry technology has provided significant advances in the creation of precise and accurate 3D models. Aerial and terrestrial photogrammetric methods have become popular choices in producing 3D models of buildings with a high level of detail. Photogrammetry has long been used for the documentation of buildings and cultural heritage objects. This technique allows the extraction of 3D information from 2D photographs, and is thus very useful in recording the architectural details of a building or structure. With the rapid development of SfM (Structure from Motion) algorithms, it has become an excellent alternative for 3D data processing (Apriansyah dan Harintaka, 2023). The results of this irradiation will obtain data called point clouds. Point clouds are a collection of 3-dimensional points that have coordinates (X, Y and Z) in the same coordinate system (Bernard Ray Barus, Yudo Prasetyo, 2017). However, despite such advancements, the semantic state of the actual building often does not fully reflect the resulting 3D model. Differences may arise in the visual representation, texture, and geometry conformity between the actual building and the model created using photogrammetric methods. In reality, complex buildings are often represented as compound buildings, which are combinations of basic primitives (Wang *et al.,* 2015; Lubis *et al.,* 2017). In order to obtain a visual impression and a topologically correct building model with semantic information, geometric constraints, such as parallel and perpendicular lines must be encoded in the building

reconstruction. Recently, there are many methods for 3D building reconstruction, which can be classified into three types: data-based methods, model-based methods, and knowledge-based methods (Wang and Yan, 2016). The photogrammetric data will then be processed using image matching algorithms and 3D reconstruction techniques to produce a 3D model of the LoD3 building. The actual semantic state of the building will be identified through field surveys and reference data collection.

Identification and classification by assigning labels and semantic attributes to objects in buildings allows us to identify and classify important elements such as walls, roofs, windows, doors, stairs, floors, and so on. This helps us understand the basic composition and structure of the building better. The method used in this study involves the collection of aerial and terrestrial photographic data. The photogrammetric data will then be processed using image matching algorithms and 3D reconstruction techniques to generate 3D models of LoD3 buildings. The actual semantic state of the building will be identified through field survey and reference data retrieval. The results of the comparative analysis will provide insight into the differences between the actual semantic state of the building and the resulting 3D model.

2. Methods

2.1 Study Area

The location of this research was carried out at the "Ratnaningsih Kinanti Universitas Gadjah Mada" women's dormitory building, Yogyakarta. This location was chosen as the object of research because it is open, not covered by vegetation or obstructed by buildings and other objects. Obstructed by buildings and other objects. Some of the problems that occurred in previous studies were due to. Some of the problems that occurred in previous studies were because the object of research used buildings that were covered by vegetation, which reduced the modeling results, so the location chosen in this study made it easier when acquiring photo data using UAVs and terrestrial photogrammetry.
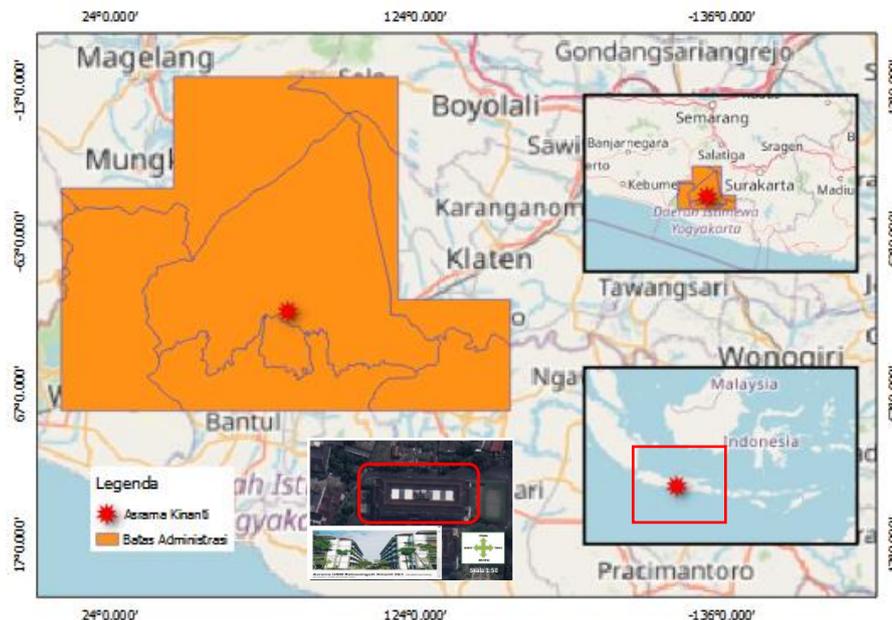

Figure 1. Map of The Study

2.2 Data Collection

Data collection was carried out for the first time, namely terrestrial photogrammetry using a type of pocket camera to acquire photo data of the first floor and second floor on the object to be modeled. Acquisition with a pocket is carried out at predetermined points by controlling the position of the camera's distance from the object to remain the same. However, there are some things that can cause the camera distance to change if the part to be acquired is blocked by other objects such as vegetation. The shooting configuration is done by going around the building until all parts of the building are acquired.

The interval of each shot is based on footsteps, this research uses 2 footsteps as a limitation of the interval of each shot which aims to make each photo taken have a good level of overlap in each photo so that during processing the object is modeled properly. From the results of terrestrial photogrammetric data acquisition, the number of photos obtained is 449 photos consisting of all parts of the building (1st and 2nd floors) that can be acquired.

Table 1. Research Data Collection

Data Type	Data Source	Output
Aerial Photo	field data	Raw
Terrestrial Photo	field data	Raw
GCP/ICP	field data	Coordinat

A DJI Phantom 4 UAV was used for data acquisition covering the entire building floor (1st, 2nd, 3rd, 4th and 5th floors). The UAV was flown at a manually set height and distance. Therefore, this

requires an experienced drone operator to avoid human error to avoid accidents during flight. In this acquisition stage, there is no flight plan path because the drone is flown manually by the operator but by still paying attention to the meeting of each photo so that when it is processed there is no noise due to the emptiness in the photo. The end lap and side lap percentages used are 80% and 65%. This value is used in consideration of the use of UAVs which are usually more unstable when compared to other mapping vehicles. The aerial photos produced at this stage were 753 photos. Shooting is done on all parts of the building from the roof to the ground. Each acquired aerial photograph has 3-dimensional coordinates or position information in the form of latitude, longitude, and altitude with World Geodetic System 84 (WGS84) as the reference ellipsoid.

2.3 Aerial Photo Data Processing and Terrestrial Photogrammetry

SfM algorithms aim to solve two types of unknowns: object structure and relative camera motion (position and orientation) in the absence of data such as the initial position of the camera and its orientation. Thus, the SfM algorithm consists of methods for feature identification and matching, homography estimation, dense reconstruction, and other methods to generate 3D coordinates of an object (Sonka dkk., 2015). Successful reconstruction with SfM requires moving the camera relative to the object. This does not necessarily mean that the camera moves around a static object, but rather to keep the camera and object at a distance to minimize errors that occur during shooting (Chiuso et al., 2000).

The ease of use of SfM with multiple sensors can enable participatory and opportunistic crowdsourced sensing opportunities, facilitating the involvement of varied data sources. Although advances in algorithms and software make the application of SfM photogrammetry quite simple in its use for topographic reconstruction, a basic knowledge of photogrammetric principles is still required for high accuracy assessment (Carbonneau & Dietrich, 2017;Eltner & Sofia, 2020). The SfM algorithm is able to generate 3-dimensional geospatial data from a series of overlapping photos in the form of points that exist in the area of the overlapping photos, which is called a dense point cloud. In addition, SfM can also determine the internal geometry and camera position and orientation automatically (Westoby et al., 2012).

In this stage, a computation is performed to reconstruct the geometry in two projection angles from the feature set and feature pairs that have been found in the previous stage (Rossi et al., 2012). Self Calibration bundle adjustment can be used to determine camera parameters such as focal length, photo center point, radial distortion, tangential distortion. By using self calibration for precise positioning and orientation, camera parameters and EOP (Exterior Orientation Parameters) are available together.

$$x_a - x_o + \Delta x = -f \cdot \frac{(r_{11}(X_A - X_L) + r_{21}(Y_A - Y_L) + r_{31}(Z_A - Z_L))}{(r_{13}(X_A - X_L) + r_{23}(Y_A - Y_L) + r_{33}(Z_A - Z_L))} \dots \dots (1)$$

$$y_a - y_o + \Delta y = -f \cdot \frac{(r_{12}(X_A - X_L) + r_{22}(Y_A - Y_L) + r_{23}(Z_A - Z_L))}{(r_{13}(X_A - X_L) + r_{23}(Y_A - Y_L) + r_{33}(Z_A - Z_L))} \dots \dots (2)$$

Note:

f = Focal Length

x_o, y_o = Image Projection center point coordinat

x_a, y_a = Lens distortion correction parameter

r = rotation matrix

X_A, Y_A, Z_A = the coordinates of the point on the object's coordinate system

X_L, Y_L, Z_L = the coordinates of the image projection center on the object coordinate system

2.4 3D Building Model Evaluation

The cityGML standard is a reference in evaluating 3D building models. CityGML is a 3D data standard published and designed by the Open Geospatial Consortium (OGC). The same object can be represented in different Levels of Detail (LOD) simultaneously based on the CityGML reference, especially for building objects consisting of 5 LODs. LOD0 is a two and a half dimensional digital terrain model. LOD1 is a block model consisting of a prismatic building with a flat roof (does not yet have a roof shape). For the LOD2 model is a building that already has a roof structure. LOD3 illustrates the form of an architectural model of a building that has a detailed wall, roof and facade structure like a typical building. LOD4 complements the previous LOD3 form by adding an interior structure. For example, the building consists of rooms, interior doors, stairs and furniture (Biljecki et al., 2016).

In addition, in evaluating the building model, geometric analysis is necessary. The geometry accuracy test analysis aims to test the level of accuracy of the building size including the position of the point or element of each building from the modeled results with the original size in the field. The original size referred to here is the measurement obtained from direct measurement on the point cloud.

The building accuracy value based on CityGML uses the Root Mean Square Error (RMSE) value reference. The RMSE value generated in the LOD 3 building model is in accordance with the CityGML standard, which is <0.5 m (Gröger, G., Kolbe, T. H., Nagel, C., & Häfele, 2008).

3. Result and Discussion

3.1 Comparative Analysis Results of Field Dimension Accuracy with Modeling

Build texture is the final stage in the 3D modeling process using the structure from motion method. At this stage is a part of the stage that aims to provide texture and color to the object so that it follows the actual state of the object.

The 3D model formed from 1201 photos and 19 pieces of GCP distribution, visually matches the model and shape of the actual object. Facades formed from modeling almost follow the original model such as doors, windows, hallways, balcony trellises can be seen in the 3D model. Similar to the point cloud, the texture produced has absolute coordinate values in the WGS84 coordinate system with the UTM Zone 49s projection system, adjusting to the coordinate system of the GCP (Ground Control Point) used during the photo processing process. texture is represented in

true color with a GSD (Ground Sampling Distance) value of 1.87 cm. With very precise GSD and appropriate color representation, texture can represent existing objects, especially building objects, very clearly and accurately, as shown.

Table 2. Comparison of the dimensions of the 3D model with the field

No	Description	Field Size (m)	Model Size (m)	Difference (m)
1	Building Width	16,76	16,82	0,06
2	Building Length	60,264	60,83	0,566
3	East Hallway	7,85	7,87	0,02
4	West Hallway	7,82	7,88	0,06
5	West Garbage Disposal Gate	1,84	1,76	0,08
6	Trellis Wall	3,582	3,53	0,052
7	South Hallway	4,68	4,75	0,07
8	North Hallway	4,75	4,83	0,08
9	Width of The Window	1,612	1,6	0,012
10	Width of Canteen Door	1,661	1,616	0,045
11	Width of WC Window	2,88	2,819	0,061
			Total	1,223236
			RMSE	11,12032727

The results of LoD3 3D modeling using point cloud data need to be tested for semantic completeness by comparing the 3D model with the original building. It can be seen that the results of the RMSE comparison between the distance of each object that is used as a sample of field measurements with a 3D model obtained an RMSE of 11.120 cm. In accordance with the thematic concept in CityGML (City Geographic Markup Language) related to model accuracy through RMSE analysis, the resulting 3D model meets the accuracy requirements in LoD3 3D modeling of <0.5m. The following shows a semantic comparison of the comprehensiveness of the 3D model of the building with its original state.

3.2 Semantic Completeness Test of 3D Model

The results of LoD3 3D modeling using point cloud data need to be tested for semantic completeness by comparing the 3D model with the original building.

Figure 2. Roof Comparison

Figure 3. South Side Building Comparison

Figure 4. North Side Building Comparison

Figure 5. East Hallway Comparison

Figure 5. West Hallway Comparison

The completeness test results can be seen that the completeness of the 3D model building elements with the original building gets accurate results. From the overall semantics of the building, when viewed from the results of the LoD3 3D model has a curvature according to the original building, where the total number of windows is 116 pieces (54 front windows and 62 rear windows, trellises with a total of 52 (26 front and 26 back, 2 doors at the front of the building and has 4 hallways spread throughout the building (north, south, west and east).

4. Conclusion

The semantic completeness test carried out by comparing the 3D LoD3 model that has been produced from the use of a combination of aerial photographs and terrestrial photogrammetry produces dimensions and semantic completeness of 3D models that are in accordance with the conditions and conditions of the original building in the field, so that the use of aerial photography and terrestrial photogrammetry in LoD3 3D modeling is assessed

from the results of the 3D modeling process automatically using SfM (Structure from Motion) technology that produces 3D building models in Level of Detail (LoD) 3 with Root Mean Square Error values <0.5 meters and has semantic completeness of buildings that match the original object based on the City Geography Markup Language (CityGML).

Comparison of the distance of each object used as a sample of field measurements with a 3D model that shows the dimensional accuracy of the building that has been modeled is obtained by 11,120 cm.

Due to the high complexity of building structures, it is still difficult to automatically reconstruct building models with accurate geometric descriptions and semantic information. To simplify this problem, this article proposes a novel approach to automatically decompose compound buildings with symmetrical roofs into semantic primitives by exploiting the local symmetry contained in the building structure. The semantic interpretation process usually involves the use of predefined guidelines or standards. For example, if there is an agreed-upon building modeling standard, semantic interpretation can refer to that standard to give consistent meanings to the elements in the 3D model.

Reference

Apriansyah, M., & Harintaka, H. (2023). Pembuatan Model 3D Bangunan LoD3 Dengan Pemanfaatan Foto Udara Dan Fotogrametri Jarak dekat. Geoid, 18(2), 243-252.

Barus, B. R., & Prasetyo, Y. (2017). Analisis Akurasi Pemodelan 3D Menggunakan Metode Close Range Photogrammetry (CRP), Unmanned Aerial Vehicle (UAV) dan Terrestrial Laser Scanner (TLS). Jurnal Geodesi Undip, 6(4), 169-179.

Biljecki, F., Ledoux, H., & Stoter, J. (2016). An improved LOD specification for 3D building models. Computers, Environment and Urban Systems, 59, 25-37.

Carbonneau, P. E., & Dietrich, J. T. (2017). Cost-effective non-metric photogrammetry from consumer-grade sUAS: implications for direct georeferencing of structure from motion photogrammetry. Earth surface processes and landforms, 42(3), 473-486.

Chiuso, A., Brockett, R., & Soatto, S. (2000). Optimal structure from motion: Local ambiguities and global estimates. International journal of computer vision, 39, 195-228.

Eltner, A., & Sofia, G. (2020). Structure from motion photogrammetric technique. In Developments in Earth surface processes (Vol. 23, pp. 1-24). Elsevier.

Gröger, G., Kolbe, T. H., Nagel, C., & Häfele, K.-H., (2008). Open Geospatial Consortium Inc . OpenGIS ® City Geography Markup Language (CityGML) Encoding Standard.

Lubis, M. Z., Anurogo, W., Oktavianto Gustin, A., Hanafi, A., Timbang, D., Rizki, F., ... & Taki, H. M. (2017). Interactive modelling of buildings in Google Earth and GIS: A 3D tool for urban planning (Tunjuk Island, Indonesia). GEOSPATIAL INFORMATION, 1(2).

Rossi, A.J., Rhody, H., Salvaggio, C., Walvoord, D.J., 2012. ABSTRACTED WORKFLOW FRAMEWORK Rochester Institute of Technology Chester F . Carlson Center for Imaging Science 54 Lomb Memorial Drive ITT Exelis Geospatial Systems 400 Initiative Drive Rochester , NY 14624. Image Process. Work. 9–12.

Sonka, 2015. Image Processing , Analysis , and Machine Vision Second Edition.

Wang, H., Zhang, W., Chen, Y., Chen, M., Yan, K., 2015. Semantic Decomposition and Reconstruction of Compound Buildings with Symmetric Roofs from LiDAR Data and. https://doi.org/10.3390/rs71013945

Wang, Q., Yan, L., 2016. A Semantic Modelling Framework-Based Method for Building Reconstruction A Semantic Modelling Framework-Based Method for Building Reconstruction from Point Clouds. https://doi.org/10.3390/rs8090737

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. Geomorphology ' Structure-from-Motion ' photogrammetry : A low-cost , effective tool for geoscience applications. Geomorphology 179, 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021