Jurnal Geografi Gea

Journal homepage: https://ejournal.upi.edu/index.php/gea

Integration Point Cloud from Aerial Photogrammetry and Action Video Camera as a Solution for 3D Mapping Technology in Urban Area

Resy Meilani[1], Tedy Imanuel Selan[1], Nida Ummatun Nadiyah[1], Siti Trisuci Putri[1], Irwan Gumilar[2], Dhota Pradipta[2], Budhy Soeksmantono[3], Agung Budi Harto[3]

[1]*Geodesy and Geomatics Engineering, ITB, Bandung, Indonesia*
[2]*Geodesy Research Division, ITB, Bandung, Indonesia*
[3]*Remote Sensing and Geospatial Information System Research Group, ITB, Bandung, Indonesia*

Correspondence e-mail: 25124019@mahasiswa.itb.ac.id

ABSTRACT

The rapid development of urban areas presents complex challenges in achieving efficient space utilization, compliance with spatial planning, and sustainable development. This study emphasizes the need for urban three-dimensional mapping to accurately represent terrestrial objects for effective city planning. By employing advanced technologies, including aerial photogrammetry and action video cameras, this study integrates point cloud data from various sources to produce high-resolution 3D spatial information. The integration process is validated through position and dimension validation methods. Position validation uses Global Navigation Satellite Systems (GNSS) with the RTK-NTRIP method and Continuously Operating Reference Stations (CORS). The Root Mean Square Error (RMSE) is calculated based on the differences in X, Y, and Z coordinates between GNSS measurements and the point cloud, resulting in a total RMSE of 0.528 meters, in accordance with the Level of Detail (LoD) 2 standards set by the Open Geospatial Consortium. Dimension validation is conducted using a distometer, comparing the measured object sizes with those in the point cloud, yielding an RMSE of 0.578 meters. These findings indicate that this integration represents efficient and accurate technology for 3D urban mapping.

ARTICLE INFO

Article History:
Submitted/Received 15 March 2025
First Revised 19 June 2025
Accepted 9 October 2025
First Available online 31 Oct 2025
Publication Date 31 Oct 2025

Keyword:
3D Urban Mapping,
Integration Mapping Technology,
Point Cloud Extraction,
Point Cloud Integration

1. INTRODUCTION

1.1 Background Research

The rapid pace of development in urban areas causes complex problems and challenges in conducting space utilization efficiency, spatial planning conformity, and sustainable development. To be able to overcome these challenges, urban three-dimensional mapping is needed that is able to represent terrestrial objects from the earth's surface in urban areas in detail and accurately so that it can be used as a reference in urban development planning (Babahajiani et al., 2017). The availability of urban three-dimensional maps in urban development planning can help stakeholders in maintaining urban infrastructure, facilitating simpler and more effective urban management, reducing redundancy, and increasing ease of communication (Šiško, 2022).

The process of providing detailed and accurate three-dimensional maps of cities, which can keep up with the highly dynamic development and complexity of urban infrastructure, requires technology that is capable of collecting spatial data in three-dimensional form quickly and accurately. The technologies currently used for three-dimensional spatial data collection are photogrammetry, laser scanners, synthetic aperture radar and handheld devices (Sirmacek and Lindenbergh, 2014; Zhu and Shahzad, 2014; Shahzad and Zhu, 2015; Tomljenovic et al., 2015). These technologies are able to generate spatial data in the form of three-dimensional points or point clouds effectively and with high accuracy that can be used for urban three-dimensional modeling (Babahajiani et al., 2017). And to handle the complexity of urban infrastructure, these technologies can be integrated to generate point clouds of urban areas as a whole, thereby improving the quality of information (Che Ku Abdullah et al., 2017).

However, current urban mapping capabilities exhibit significant gaps. Traditional 2D maps struggle to predict complex multi-level infrastructure conditions. Despite technological advancements, there are still many limitations in current 3D mapping approaches. Aerial methods using UAVs achieve excellent horizontal accuracy but fail to capture crucial vertical surfaces and street-level infrastructure hidden beneath overhangs (Nex, F and Remondino, 2014). Conversely, terrestrial laser scanning provides detailed facades but requires expensive equipment and time-consuming operations, making it impractical for rapid large-scale mapping.

This research gap becomes evident when examining integration attempts. While (Persad and Armenakis, 2017) successfully integrated professional UAV and mobile LiDAR data, their approach requires equipment costing hundreds of thousands of dollars, limiting accessibility for developing nations. (Gonçalves and Pinhal, 2018) demonstrated that consumer action cameras could achieve 0.5-1.0 m accuracy, but no comprehensive workflow exists for integrating such affordable terrestrial data with aerial photogrammetry. This gap is critical, as developing countries urgently need cost-effective solutions for urban planning, disaster management, and informal settlement mapping.

Rapid mapping technology can utilize aerial photogrammetry techniques and video recording with action video cameras mounted on moving vehicles (Gonçalves and Pinhal, 2018). Aerial photogrammetry has high resolution and advantages in terms of large area coverage. Meanwhile, action cameras mounted on moving vehicles are a fast, cheap, and accurate terrestrial mapping solution in urban areas compared to other terrestrial three-dimensional mapping technologies such as laser scanners, and are capable of extracting point cloud data using photogrammetric principles (Gonçalves and Pinhal, 2018). Point clouds generated from both technologies can be integrated using data integration algorithms such

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

as Target-to-Target Method (Li, D et al., 2022), Random Sample Consensus (RANSAC) (Gu, X et al., 2020), and Iterative Closest Point (ICP) (Che Ku Abdullah et al., 2017; Gu, X et al., 2020) to obtain complete urban area point cloud data.

This research focuses on integrating point clouds from videogrammetry using action cameras with point clouds generated from aerial photographs using UAVs. Previous studies have applied videogrammetry technology in 3D mapping of historical buildings, such as the work by (Pepe et al., 2022), which utilized a Ricoh Theta Z1 spherical camera to capture 360-degree video and processed it with Structure from Motion (SfM) and Multi-View Stereo (MVS) algorithms. The results demonstrated rapid data acquisition and broad area coverage; however, challenges remain regarding accuracy and point cloud quality. Integrating UAV data is expected to enhance the quality and detail of 3D modeling, addressing the limitations present in traditional videogrammetry methods.

Moreover, (Nuryudha, 2024) conducted a comparative analysis of 3D modeling on the Gunung Gangsir Temple using a DJI Phantom 4 Pro V2 UAV and a Canon 7D DSLR camera. This method relied on data processing through the SfM algorithm to generate a 3D model from captured photos and videos. The findings indicated that photogrammetry produced models with better detail and sharpness, achieving 20,494,298 points compared to 12,710,860 points from videogrammetry. While photogrammetry demonstrated lower RMSE and higher accuracy, videogrammetry offered faster data acquisition and processing. These findings highlight the strengths and weaknesses of each method in documenting cultural heritage.

Therefore, in this research, the integration process of point cloud data generated from aerial photogrammetry and action video cameras will be carried out to obtain complete three-dimensional spatial data. The aerial photogrammetry used is derived from the results of UAV data acquisition and the action video camera used is GoPro Hero 9. The integration of point clouds from videogrammetry and UAV data in this study aims to leverage the speed of videogrammetry while enhancing accuracy with UAV imagery. This approach seeks to provide a comprehensive solution for urban mapping and 3D modeling of diverse urban objects. The results of the data processing process will be tested for accuracy based on the standards set by the IFC BIM Forum in 2012 to determine the feasibility of the data.

1.2 Theoretical Framework

1.2.1 Fundamental Concepts of Point Cloud from Photogrammetry and Videogrammetry

Point clouds represent digital surface geometry as a collection of three-dimensional points with X, Y, and Z coordinates. Recent developments in point cloud technology have transformed urban mapping, enabling applications such as Building Information Modeling (BIM) and large-scale mapping using aerial photos or LiDAR (Xiao et al., 2023).

One fundamental principle in photogrammetry is the collinearity condition (**Figure 1.)**, which describes the relationship between a point object in the real world, the camera's projection center, and the point on the photo plane; these three elements must lie on a straight line to ensure accurate mapping. Once data is obtained, the next step is to transform coordinates from the photo system to the ground coordinate system, involving translation, rotation, and scale changes. Camera calibration is also crucial for correcting internal orientation parameters (IOP), such as focal length and lens distortion coefficients, thereby improving the accuracy of 3D modeling. Radial and tangential distortions can affect image quality, where radial distortion is the linear shift of photo points relative to the principal point, and tangential distortion arises from centering errors in lens elements.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

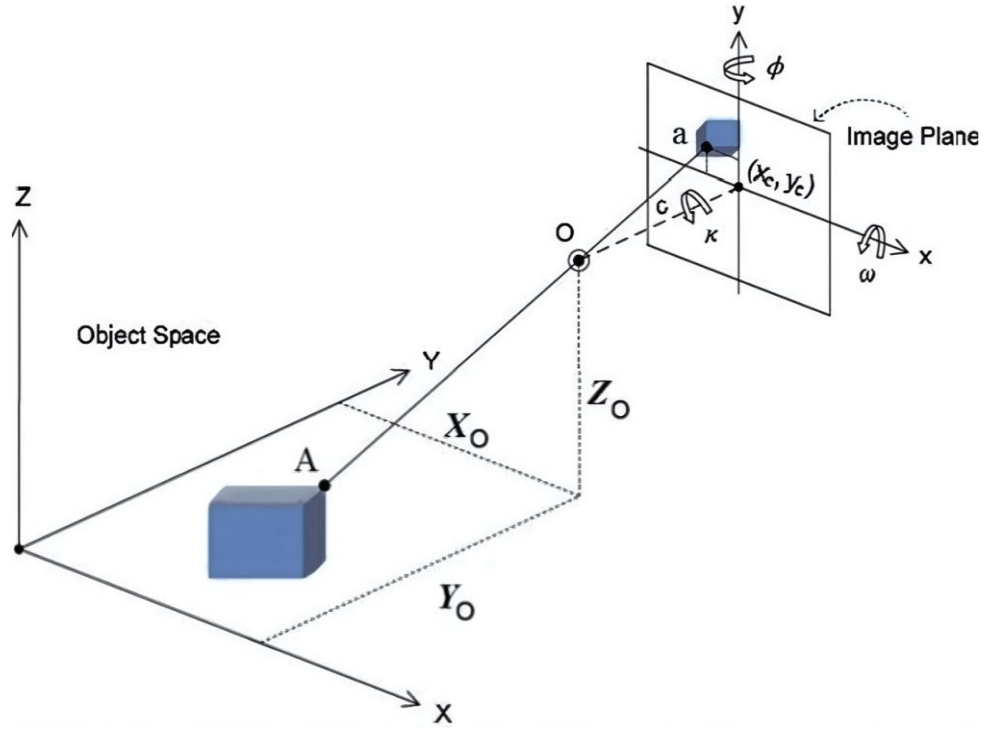

Figure 1. Collinearity condition in Photogrammetry adapted from (Situmorang, 2019)

The terrestrial collinearity equations (1) in the case of photogrammetry is as follows.

$$x - xo = -c[\frac{(m_{11}(X - X_o) + m_{12}(Z - Z_o) + m_{13}(Y_o - Y))}{(m^{31}(X - X_o) + m^{32}(Z - Z_o) + m^{33}(Y_o - Y))}]$$

$$y - yo = -c[\frac{(m^{21}(X - X_o) + m^{22}(Z - Z_o) + m^{23}(Y_o - Y))}{(m_{31}(X - X_o) + m_{32}(Z - Z_o) + m_{33}(Y_o - Y))}] \quad (1)$$

Description:

x, y: The image point coordinates.
X, Y, Z: The object point coordinates.
X_o, Y_o, Z_o: The coordinates of the exposure station in the object space system.
c: The focal length of the camera.
m_{ij}: The elements of the rotation matrix.
x_o, y_o: The coordinates of the principal point.

In the generation of point clouds, photogrammetric principles play a vital role. This process relies on the collinearity principle that connects image coordinates to ground coordinates (Situmorang, 2019). Modern Structure from Motion (SfM) algorithms **(Figure 2.)** have automated this process, with recent implementations achieving sub-pixel accuracy through deep learning-enhanced feature matching (Eltner and Sofia, 2020).

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

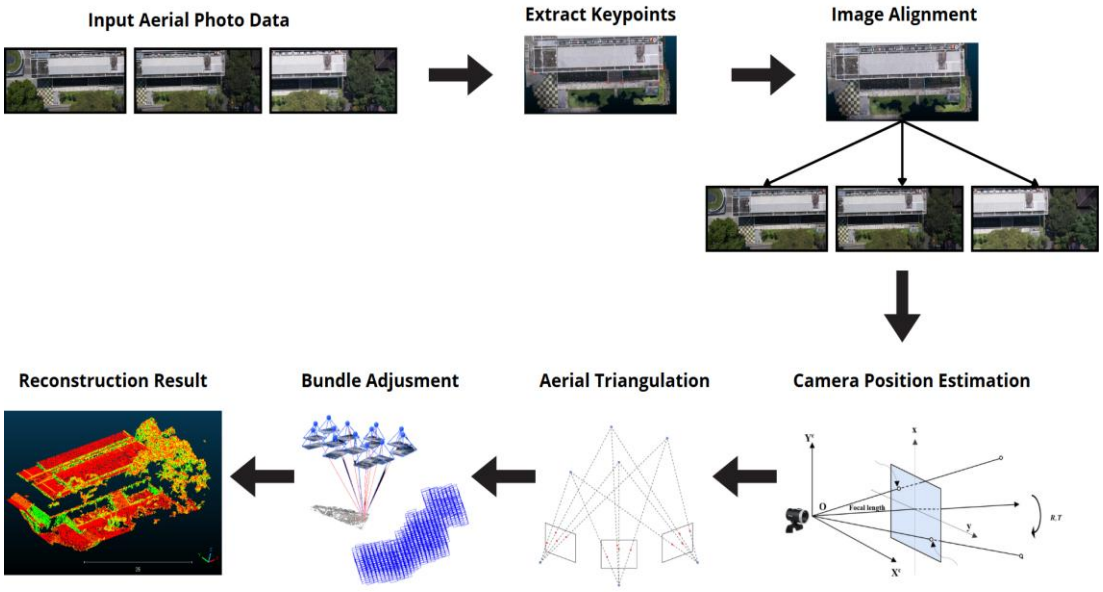

Figure 2. Process Structure of Motion adapted from (Eltner and Sofia, 2020)

Videogrammetry is a three-dimensional mapping technique that uses video recordings to generate 3D models, unlike photogrammetry, which relies on still images. The main advantage of videogrammetry lies in its ability to capture moving objects and generate data quickly, allowing for real-time analysis in dynamic environments, such as urban areas. While photogrammetry uses the principle of collinearity to connect image points and object points from different images, videogrammetry leverages feature matching between video frames, thus enhancing flexibility and efficiency in data acquisition.

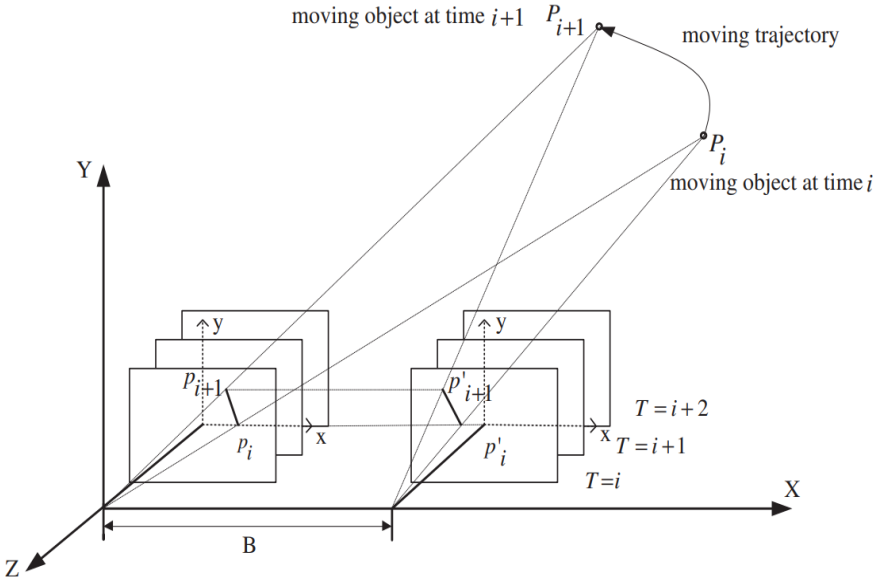

Figure 3. Principle of videogrammetry

Figure 3. displayed explains the concepts of calculating instantaneous velocity and the centroid of a moving object in the context of 3D coordinates. Instantaneous velocity is determined by analyzing the changes in the coordinates of a feature point at various time

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

stages, allowing for the assessment of how fast the object moves in each direction (X, Y, Z). Additionally, the centroid of the moving object is calculated based on the 3D coordinates of all feature points, providing an average position that is useful for analyzing the object's motion. Both concepts are highly relevant in videogrammetry, where tracking moving objects through video recordings requires an accurate understanding of the dynamics of motion and the position of objects in space.

1.2.2 Theory of Multi-Source Point Cloud Integration

In point cloud integration research, various methods can be employed, each with its own advantages and limitations. Feature-based methods, for instance, require feature extraction that is often challenging on dense point cloud data, such as from UAV aerial imagery, or unstructured data like video from action cameras. On the other hand, global integration methods like Fast Global Registration, while offering better robustness against initial conditions, often require longer computational times. These considerations are crucial when selecting the appropriate algorithm for integration (Qian-Yi Zhou et al., 2016).

Based on a comprehensive literature review, the Iterative Closest Point (ICP) method was chosen as the integration algorithm for this research. ICP is a standard algorithm that has proven effective for integrating point clouds from various data sources. The characteristics of the data in this study align with the fundamental assumptions of ICP, where both point clouds have a sufficiently good initial alignment following the georeferencing process. Assuming that the data from UAVs and action camera videos share the same scale, the rigid integration approach, which involves only rotation and translation without scaling changes, becomes highly relevant. The choice of ICP is based on its ability to address the limitations of other methods, making it an optimal choice for achieving accurate and efficient integration results.

The Iterative Closest Point (ICP) algorithm is a widely used method for aligning two point cloud datasets, known as the source point cloud (P) and the target point cloud (Q). The process begins by establishing initial correspondences between the two clouds, where each point in Q is matched with the nearest point in P. Subsequently, the algorithm calculates the optimal transformation parameters, consisting of a rotation matrix (R) and a translation vector (t), to minimize the distance between corresponding points. This transformation, represented by the equation:

$$P' = s.(R.Q) + t \qquad (2)$$

is applied to Q, where P' is the transformed point cloud (output), s is the scale factor (with s=1), R is the rotation matrix, Q is the reference point cloud, and t is the translation vector. This transformation is applied to Q, after which the registration error is computed. The algorithm iterates through these steps until the error is minimized below a predefined threshold or a maximum number of iterations is reached.

2. METHODS

2.1 Study Area

The study area in this research is the Campus Center Timur building of Bandung Institute of Technology located in Bandung City, as shown in **Figure 4.** This building was chosen because it is considered to represent an urban building object that has diverse architecture and details. The building is equipped with a colonnade structure and prominent staircase details.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

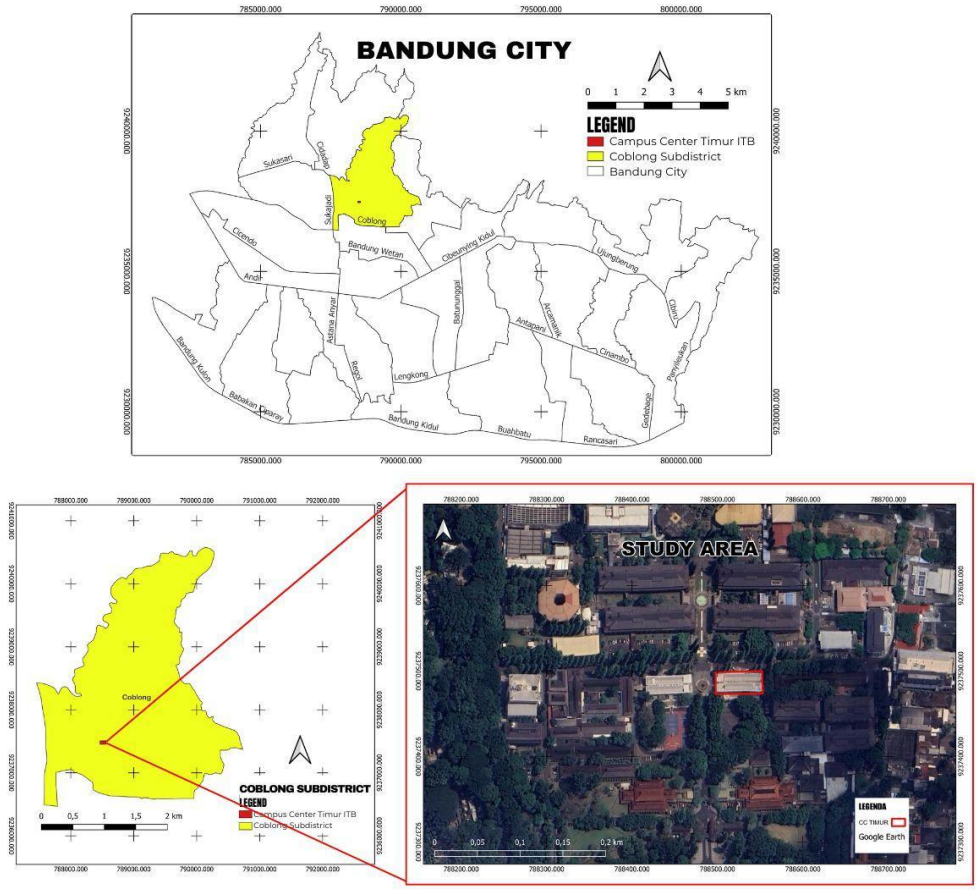

Figure 4. Study Area in Campus Center Timur ITB, Bandung City

Figure 5. shows details of the Campus Center Timur ITB object in Bandung City from various angles that provide a comprehensive overview of the building. This representation from different angles provides a clearer view of the building's shape and design.

Figure 5. Object details of Campus Center Timur ITB, Bandung City

2.2 Data

The data used in this study is presented in **Table 1.** This table includes different types of data obtained from different sources. The data is the result of acquisition or primary data taken in 2024.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

Table 1. Research Data

Data type	Source data	Data density	Distance between points	Coordinate system	Vertical datum	Form at
Point cloud	Aerial photogrammetry	291 pts/m^2	4-6 cm	UTM Zone 48s	Ellipsoid	.las
	Action video camera	676 pts/m^2	1 cm	UTM Zone 48s	Ellipsoid	.las

2.3 Methodology

Figure 6. is a general workflow for the integration of point clouds from aerial photogrammetry and action video cameras as a solution for 3D mapping technology in urban areas. This diagram shows how point cloud data is generated from two sources: aerial images captured by an Unmanned Aerial Vehicle (UAV) equipped with an ILCE-6000 camera, and video recorded using a GoPro Hero 9. The process begins with the UAV capturing photos, while data collection with the GoPro Hero 9 involves recording video and GPS data in the modeling area. Each process utilizes image matching principles, resulting in separate point clouds. These two point clouds are then integrated, followed by a validation process to ensure the accuracy of positional and distance dimensions.

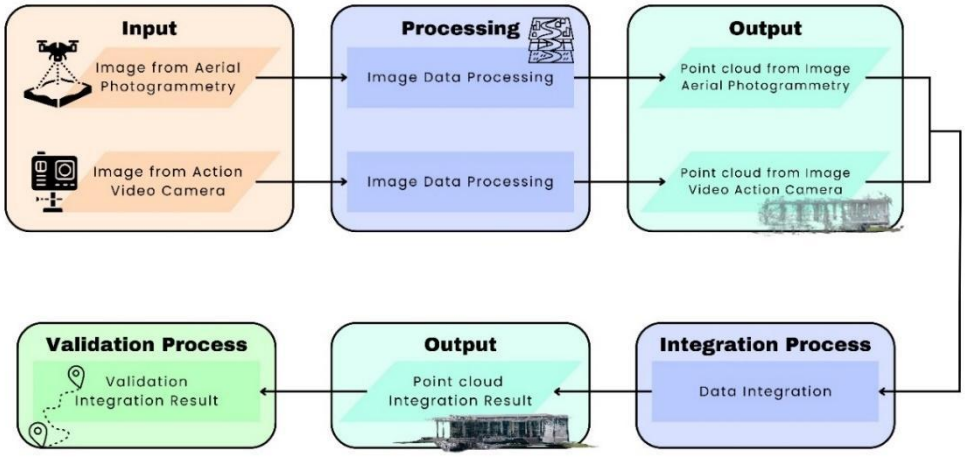

Figure 6. General Research Workflow

2.3.1 Methodology of Point Cloud Extraction from Aerial Photogrammetry

Digital photogrammetry produces spatial data in the form of orthophoto maps and three-dimensional models by processing overlapping aerial images. This method enables the reconstruction of real-world objects with high geometric fidelity and accuracy (Ihsan and Sugandi, 2019). The process of extracting point clouds from aerial photogrammetry begins with capturing photos using an Unmanned Aerial Vehicle (UAV) equipped with an ILCE-6000 camera. A total of 1,369 images were taken at an altitude of 97.5 meters, with a ground resolution of 1.47 cm per pixel, and overlaps between images reaching 80% with a sidelap of 60% to ensure modeling accuracy. This configuration follows common photogrammetric principles for high-accuracy modeling, in line with procedures outlined in SNI 9135-1:2022, which emphasize proper planning of flight paths and GCP distribution to meet large-scale mapping requirements (Kresna Sumaamijaya and Gilang Attalla, 2025).This data acquisition covers a total area of 0.603 km². However, this study focuses solely on the CC Timur building at ITB. Next, camera calibration is performed to correct lens distortion that affects the results.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

The camera calibration coefficients indicate a measured focal length (F) of 6506.49 with a small error of 0.11, indicating good accuracy. The principal points (Cx and Cy) are -28.674 and -29.422, with errors of 0.037 and 0.029, respectively. The radial distortion coefficients K1, K2, and K3 are -0.0336147, -0.0625118, and 0.0000000, respectively, indicating low radial distortion. The asymmetric distortion coefficients P1 and P2 are 1.58081e-05 and 0.000701747. This calibration demonstrates that the camera has been well-calibrated for subsequent geometric processing.

In photogrammetric processing, the imaging model uses the principle of collinearity, involving external orientation parameters (position and rotation of images) as well as internal orientation of the camera (principal point and focal length). This model also considers image distortions, including radial, asymmetric, and tangential distortions. After camera calibration, processing begins with feature matching to identify and match points on the same object between images. The aerial triangulation process is performed to increase control points photogrammetrically, with control points evenly distributed across the acquisition area. This process successfully identified 438,950 tie points from a total of 1,938,786 proposed points, achieving an acceptance rate of 22.6%. The quality of the tie points is very good, with an average residual of 0.67 pixels and an overall RMSE of 0.78 pixels. This average residual indicates how accurately the positions of the tie points are represented in pixels, while the overall RMSE measures the average error of all generated points. The generation of point clouds is achieved through image matching processing (**Figure 7.**), which utilizes the matched features from overlapping images to reconstruct the 3D structure of the scene. The process ensures position accuracy through the use of control points measured with high precision using GNSS. These control points serve as references to improve the positions of points generated from aerial photographs. During the bundle adjustment process, data from the control points is used to adjust the positions and orientations of the points. After adjustment, the RMSE is calculated, resulting in a value of 19.0732 cm. This value indicates how closely the positions of points in the point cloud align with the accurately measured values, reflecting the accuracy of the final results. Points generated from feature matching between images are also examined and adjusted based on information from the control points, thereby enhancing overall accuracy.

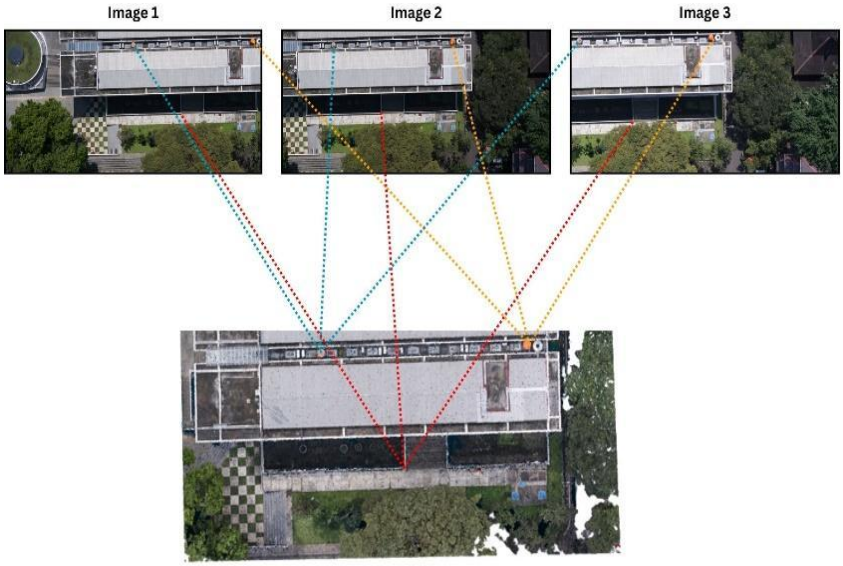

Figure 7. Example Visualization of How Point Clouds are Generated from Aerial Photogrammetry Images

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

2.3.2 Methodology of Point Cloud Extraction from Action Video Camera

The point cloud extraction process from the action video camera is carried out using a photogrammetric pipeline model, starting from data collection, processing, to producing output. The flowchart of the entire point cloud extraction process from the action video camera is shown in **Figure 8.**

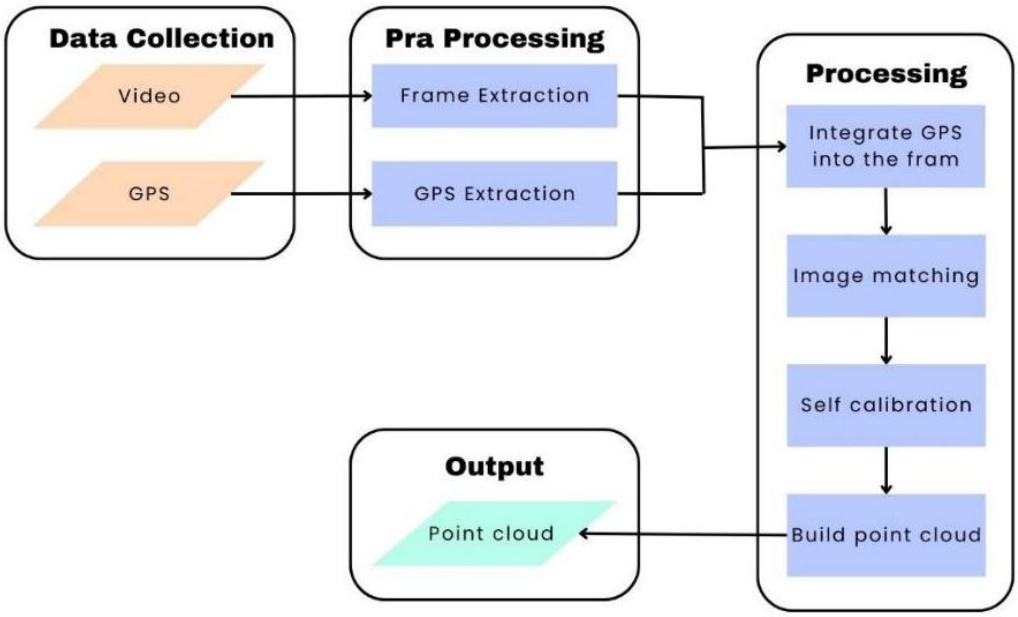

Figure 8. Workflow of Point Cloud Extraction from Action video camera

Data collection was conducted using a GoPro Hero 9 Black action video camera mounted on a vehicle, which features an integrated GNSS receiver. The GPS specifications include a positional accuracy of ±3-5 meters, typical for consumer-grade devices, an update rate of 1 Hz (one position fix per second), and output coordinates in the WGS84 format (latitude, longitude, and ellipsoidal height). This camera was selected for its speed, lightweight design, durability, and ability to operate in challenging conditions, making it suitable for photogrammetric data acquisition (Gonçalves and Pinhal, 2018). Data collection involved recording video and GPS in the area to be modeled. Videos were recorded at 1080p resolution with a frame size of 1920 × 1080 pixels, a video recording rate of 50 frames per second (FPS), and a linear processing mode, which helps reduce rolling shutter effects and radial distortion (Gonçalves and Pinhal, 2018). Each video frame is linked to GPS coordinates through metadata timestamps, allowing precise alignment, while coordinates between GPS fixes are interpolated using the vehicle's speed and direction to maintain spatial continuity. Transformation of GPS coordinates to UTM Zone 48S is also performed to ensure compatibility with the photogrammetry coordinate system.

The recording results will produce video files in .MP4 format, so at the preprocessing stage, it is necessary to extract the frames to obtain images in JPEG format and GPS positions from each video frame. Frame extraction is performed to separate the video frames, while GPS positions are extracted from the video metadata. This metadata provides information on the recording time of each video frame, as well as latitude and longitude coordinates, height above the ellipsoid, and speed for each frame. The extraction results are saved in .csv format to be used as input data for defining the position of each extracted frame.

Based on the results of frame extraction and GPS positions, data processing is carried out to obtain a point cloud from the video recording. This process begins by entering the

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

extracted GPS position data to define the position of each frame. Then matching between photos is carried out based on the results of keypoint calculation and 3D model reconstruction using Automatic Aerial Triangulation and Bundle Block Adjustment techniques (Hapriansyah and Hidayat, 2021). During the bundle block adjustment process, a self-calibration process is performed to optimize the internal and external parameters of the camera based on the keypoint matching results (Suziedelyte Visockiene et al., 2014). Since in this data processing process the GPS position of each frame is defined, the camera position of each frame is readjust based on the 3D model reconstruction results (Hapriansyah, 2021).

The GPS and video synchronization method uses a systematic approach to ensure accurate spatial-temporal matching. Each video frame is connected to GPS coordinates through timestamps in the metadata, allowing for precise alignment. For frames captured between GPS fixes, coordinates are interpolated based on the vehicle's speed and direction to maintain spatial continuity. GPS coordinates are also transformed from the WGS84 system to UTM Zone 48S for compatibility with the photogrammetry coordinate system.

In initial tests, the observed positional variations between consecutive frames were ±2-5 meters, and altitude changes were ±3-8 meters, which are typical values for consumer-grade GPS receivers. To calculate the time for each video frame, the following interpolation formula (3):

$$t_n = t_1 + \frac{n-1}{50} \qquad (3)$$

Where t_1 is the time of the first frame and n is the frame number. With a speed of 50 fps, this formula allows for precise timing of each frame captured. To improve position accuracy, a camera position adjustment method is applied through iterative photogrammetric optimization using Bundle Block Adjustment. The azimuth of the vehicle's trajectory is calculated using the formula (4):

$$\alpha_i = tan^{-1}\left(\frac{N_{i+1} - N_{i-1}}{E_{i+1} - E_{i-1}}\right) \qquad (4)$$

Where E and N are planar coordinates obtained from the GPS data. This azimuth is crucial to ensure that the camera's direction aligns with the vehicle's trajectory, allowing for accurate calculation of observed object positions, such as traffic signs. The coordinates of objects in the image can be calculated by combining the azimuth and camera position formula (5):

$$x_{object} = tan^{-1}\left(\frac{f}{x_{object}}\right) + \alpha_k + \beta$$

Where f is the focal length, x_{object} is the coordinate in the image, and β is the angle between the vehicle's axis and the trajectory.

The 3D model reconstruction process that has been previously carried out will produce tie points. Based on these results, the point cloud extraction process is performed using the principle of image matching (**Figure 9.**). In this research, point cloud extraction is conducted at the optimal density option, which means that every four pixels in the image will be used in the point cloud formation process. The results of the point cloud extraction from all images will be automatically saved in .las format.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

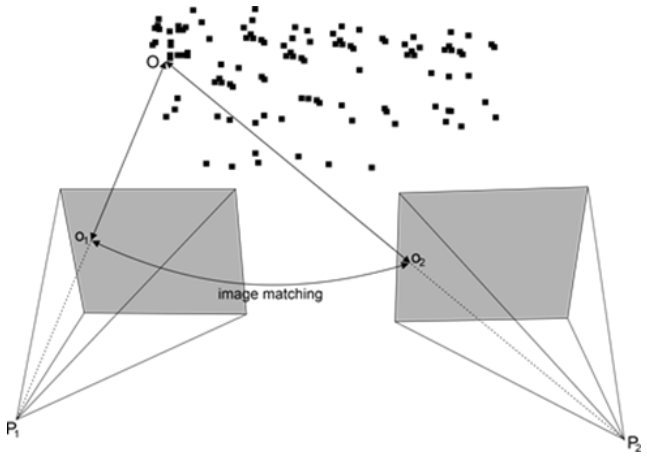

Figure 9. Illustration of Image matching

This adjustment successfully improved overall position accuracy by 60-75%, as well as increased frame connectivity from 65% to 92%. With this approach, the data collected from the GoPro Hero 9 in the study area is well-integrated and accurate, supporting the reliability of the resulting mapping.

2.3.3 Methodology of Point Cloud Integration between Aerial Photogrammetry and Action Video Camera

Figure **10.** illustrates the workflow for integrating point clouds between aerial photogrammetry and action video cameras. The process begins with gathering data from both sources, followed by a processing stage that includes filtering and coarse registration. After that, fine registration is performed using the Iterative Closest Point (ICP) method to enhance alignment accuracy. The integration results are then validated for both position and dimension.

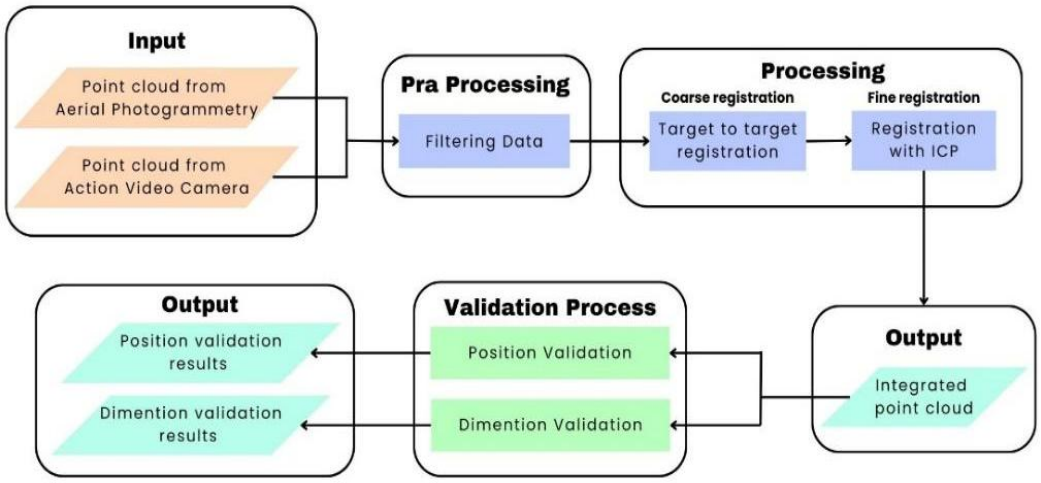

Figure 10. Workflow of Point Cloud Integration between Aerial Photogrammetry and Action Video Camera

The integration process of point clouds from aerial photogrammetry and action video cameras begins with the input stage, which involves using point cloud data extracted from each technology source. Next, a pre-processing stage is carried out, which includes data filtering to enhance data quality by removing irrelevant point clouds. Following this, the processing stage encompasses two main sub-stages: Coarse Registration and Fine Registration. In Coarse Registration, Target-to-Target registration is performed to roughly

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

align the two point clouds. A minimum of four pairs of points from both datasets are required in this process to generate the transformation matrix for translation, rotation, and scale (Yang et al., 2021). After that, Cloud-to-Cloud registration is conducted using Iterative Closest Point (ICP) in the Fine Registration stage to optimize the previous translation, rotation, and scale transformations, resulting in a more accurate point cloud integration (Li, P et al., 2020).

2.3.4 Methodology of Validation

Once the integration process is complete, a validation process for the position and dimensions of the point cloud is conducted to assess the integration results. Position validation is performed using Global Navigation Satellite Systems (GNSS) with the Real Time Kinematic-Networked Transport of RTCM via Internet Protocol (RTK-NTRIP) method. In this measurement, a base station known as a Continuously Operating Reference Station (CORS) is used. The CORS used as the basis for GNSS measurements in this validation is CLBK in Lembang, owned by the Geospatial Information Agency (BIG), with a maximum baseline length of 7.5 kilometers. The Root Mean Square Error ($RMSE$) is then calculated as a quality parameter by using the differences between the X, Y, and Z coordinates obtained from the GNSS measurements and the point cloud at all sample points.

The position validation $RMSE$ equation can be seen in the following equation where n is the number of validation samples.

$$RMSE_{each\ point} = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \quad (1)$$

$$RMSE_X = \sqrt{\sum_{i=1}^{n} \frac{\Delta x_i}{n}} \quad (2)$$

$$RMSE_Y = \sqrt{\sum_{i=1}^{n} \frac{\Delta y_i}{n}} \quad (3)$$

$$RMSE_Z = \sqrt{\sum_{i=1}^{n} \frac{\Delta z_i}{n}} \quad (4)$$

$$RMSE_{total} = \sqrt{(RMSE_X)^2 + (RMSE_Y)^2 + (RMSE_Z)^2} \quad (5)$$

Dimension validation is carried out using a distometer. The length of the object's dimension measured by the distometer is then compared with the dimension length in the point cloud, resulting in an $RMSE$ value. The equation for the $RMSE$ of dimension validation can be seen in the following equation, where nnn is the number of validation samples.

$$RMSE_{Dimention\ total} = \sqrt{\sum_{i=1}^{n} \frac{\Delta l_i}{n}} \quad (6)$$

3. RESULTS AND DISCUSSION

3.1 Point Cloud Extraction from Aerial Photogrammetry

The point cloud data generated from the image matching process of aerial photos results in 438,950 tie points from a total of 1,938,786 proposed points. After all accuracy requirements for aerial triangulation are met, a dense point cloud is produced with a total of 258,178,797 points (**Figure 11.**), complete with color attributes.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

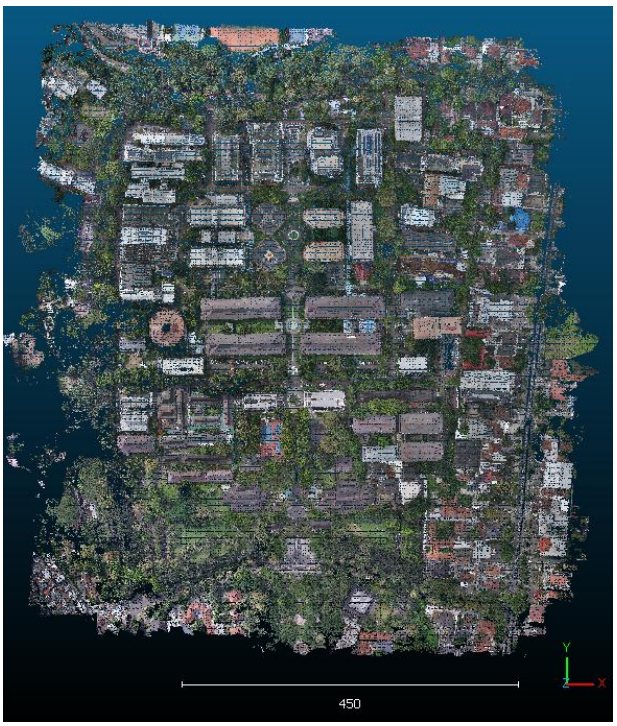

Figure 11. Point cloud from Aerial Photogrammetry Result

In the focus area of the study, the data was cropped to encompass only a specific part of the building and its surroundings, resulting in a total of 2,331,639 points (**Figure 12.**). This crop aims to analyze in detail the structure of the building that is the subject of the research, as well as to enhance efficiency in the subsequent data processing.

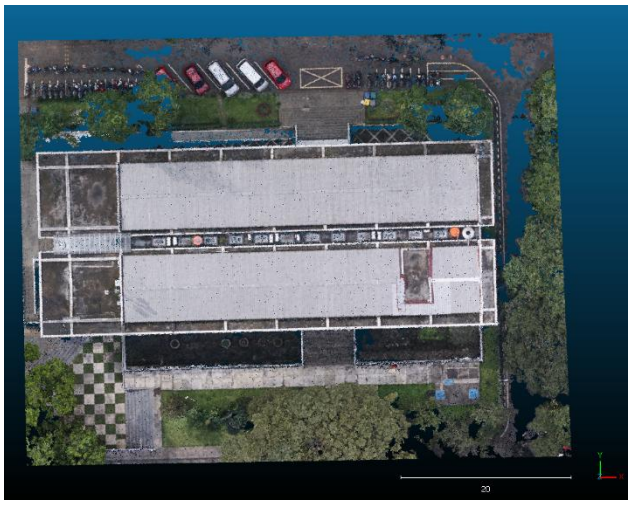

Figure 12. Point cloud from Aerial Photogrammetry Result in Area Study

As part of the accuracy evaluation of point cloud data acquisition, Table 2. and Table 3. below present the average errors in camera location and orientation. This data provides insights into the accuracy of the results obtained from aerial photogrammetry, as well as indicating the potential impact of these errors on the quality and reliability of the generated point cloud. Table 2. and Table 3. show that the total positional error is 14.9147 cm. On the other hand, the total camera orientation error of 0.69812° indicates that the camera orientation is relatively accurate, which could positively contribute to the quality of the generated point cloud.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

Table 2. Average camera location error

X error (cm)	Y error (cm)	Z error (cm)	XY error (cm)	Total error (cm)
13.2625	5.64208	3.83672	14.4127	14.9147

Table 3. Average camera orientation error

Omega error (°)	Phi error (°)	Kappa error (°)	Total error (°)
0.101117	0.142191	0.675964	0.69812

Table 4. show the Root Mean Square Error (RMSE) for various control points, reflecting the accuracy of the mapping. Based on the total error value, a RMSE of 19.0732 cm was obtained, indicating the accuracy of the point cloud generated from aerial photogrammetry.

Table 4. Check point RMSE

Point Id	X error (cm)	Y error (cm)	Z error (cm)	Total (cm)	Image (pix)
GCP01	-4.58453	13.9676	16.3662	21.9991	0.491 (8)
GCP08	1.66516	0.590027	21.3122	21.3853	0.519 (9)
GCP09	-10.2065	-1.99274	18.8558	21.5333	0.608 (5)
GCP10	5.17616	0.969172	16.3667	17.193	0.411 (7)
GCP24	-7.69272	-4.555	-6.19949	10.8793	1.328 (8)
Total	**6.54115**	**6.6498**	**16.6367**	**19.0732**	**0.762**

3.2 Point Cloud Extraction from Action Video Camera

The point cloud generated from the action video camera as shown in Figure 13.

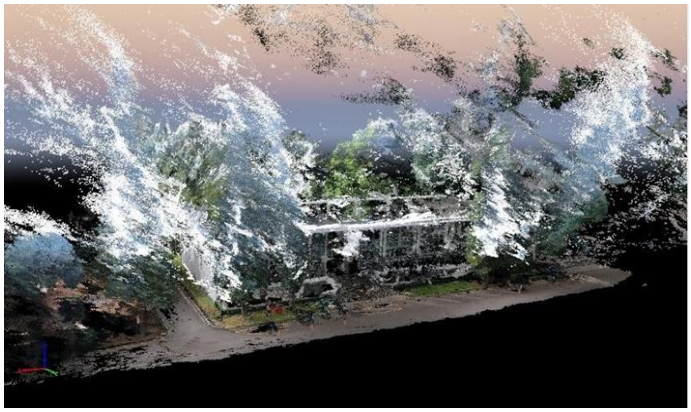

Figure 13. Point Cloud from Action video camera

The results of point cloud processing heavily depend on the quality of the 3D model reconstruction process and the accuracy of the GPS from the action video camera. Semantically, the point cloud generated by the action video camera can effectively capture the details of objects because they are recorded at a close distance to the camera; however, the noise present in the point cloud is also quite significant. Spatially, the accuracy of the point cloud is greatly influenced by the GPS accuracy of the action video camera and the readjust process conducted during 3D model reconstruction. The extent of positional change during the readjust process is shown by the absolute geolocation variation values, as presented in **Table 5.**

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

Table 5. Absolute geolocation variance

Min Error (m)	Max Error (m)	Geolocation Error X (%)	Geolocation Error Y (%)	Geolocation Error Z (%)
-	-15.00	0.00	0.00	0.00
-15.00	-12.00	0.00	0.00	0.00
-12.00	-9.00	0.00	0.00	0.00
-9.00	-6.00	0.00	0.00	0.00
-6.00	-3.00	0.00	0.00	0.00
-3.00	0.00	46.28	56.76	53.19
0.00	3.00	53.72	43.24	46.81
3.00	6.00	0.00	0.00	0.00
6.00	9.00	0.00	0.00	0.00
9.00	12.00	0.00	0.00	0.00
12.00	15.00	0.00	0.00	0.00
15.00	-	0.00	0.00	0.00
Mean (m)		0.000008	-0.000241	-0.000888
Sigma (m)		0.487970	0.393635	0.707503
RMS Error (m)		0.487970	0.393635	0.707503

The minimum and maximum error values for absolute geolocation variation represent the geolocation error interval between -1.5 and 1.5 times the maximum accuracy across all images. The geolocation error columns for the X, Y, and Z axes indicate the percentage of images with geolocation errors that fall within the specified error intervals. The results show that data from the action video camera exhibit geolocation error variance within the range of -3.00 to +3.00 on each axis, with the RMS error values for each axis being 0.487970 meters for the X axis, 0.393635 meters for the Y axis, and 0.707503 meters for the Z axis. The magnitude of this absolute geolocation error may be due to inaccuracies in the GPS of the action video camera. This can be observed in the visualization of GPS position changes for each frame before and after readjust, as illustrated in **Figure 14**.

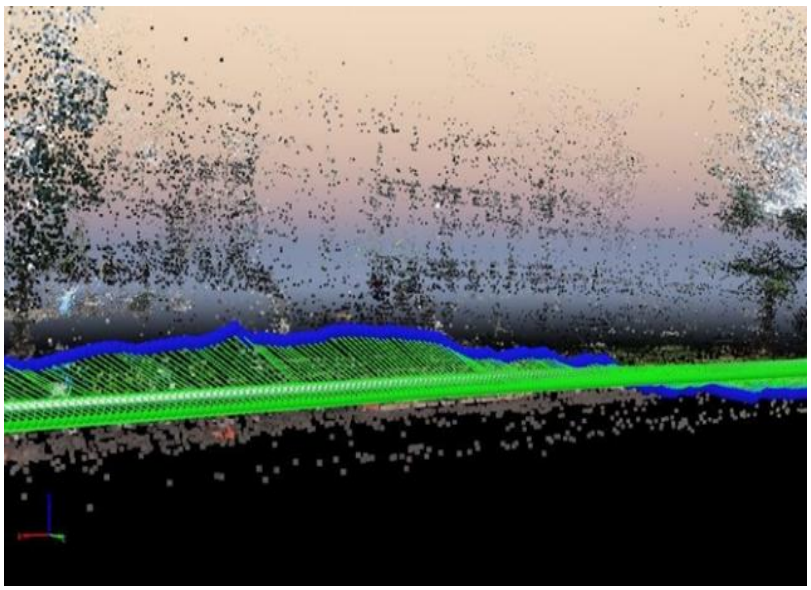

Figure 14. Readjust result

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

The GPS position before readjust (blue points) shows rapid variations in position changes in each frame, while after readjust (green points), the position changes during the 3D model reconstruction process tend to be more stable. Therefore, in this study, the spatial accuracy of the point cloud extraction results from the action video camera will be improved during the data integration process with the point cloud from stereo images.

3.3 Point Cloud Integration between Aerial Photogrammetry and Action Video Camera

The point cloud results from the aerial UAV imagery and action video camera before and after the filtering process can be seen in Figure 15. and Figure 16. The filtering process was carried out manually to eliminate point clouds that were outside the main objects as well as dynamic objects, such as trees and vehicles in the parking lot. This process was conducted to achieve a high-quality integration of the point clouds from the aerial UAV imagery and the action video camera.

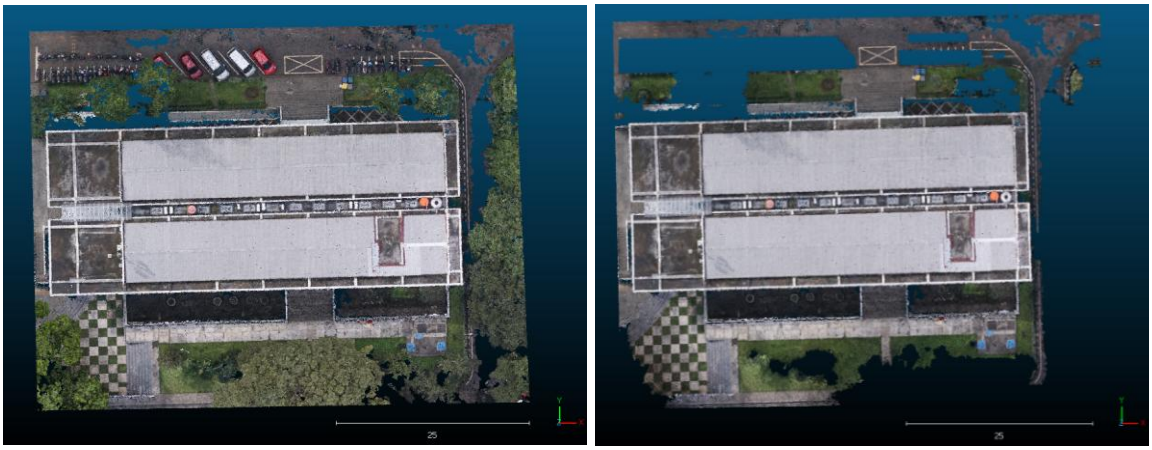

Figure 15. Aerial photogrammetry point cloud results before (left) and after (right) filtering process

Figure 16. Action video camera point cloud results before (left) and after (right) filtering process

After the filtering process, a coarse integration of the point cloud was performed using the target-to-target method. In this integration process, a total of five points were used, as illustrated in Figure 17.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

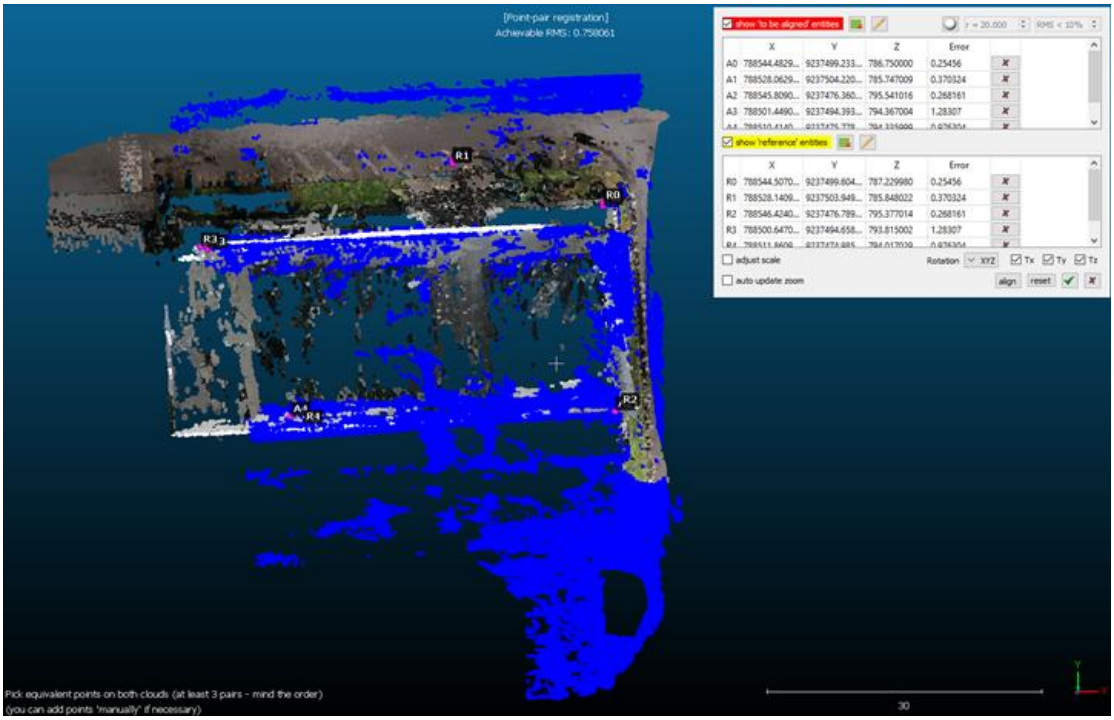

Figure 17. The points used in the coarse registration process

Subsequently, a fine registration process was conducted using the cloud-to-cloud method. The overlap between the point cloud from the images and the action camera was nearly 50%, thereby meeting the requirements for the cloud-to-cloud integration method, which is set at 30% to 40%. The quality of this integration process is assessed by the RMSE value, which is 0.758061 meters, as shown in **Figure 18**.

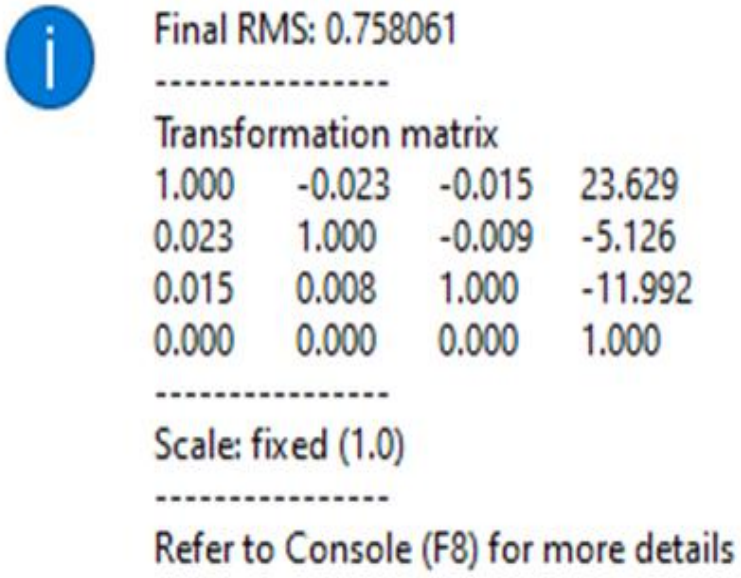

Figure 18. Result of Fine Registration

The results of the point cloud integration from the images and action camera can be viewed from the top, north-south, and east-west perspectives in Figure 19. until **Figure 23**.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

Figure 19. Top view of point cloud integration result

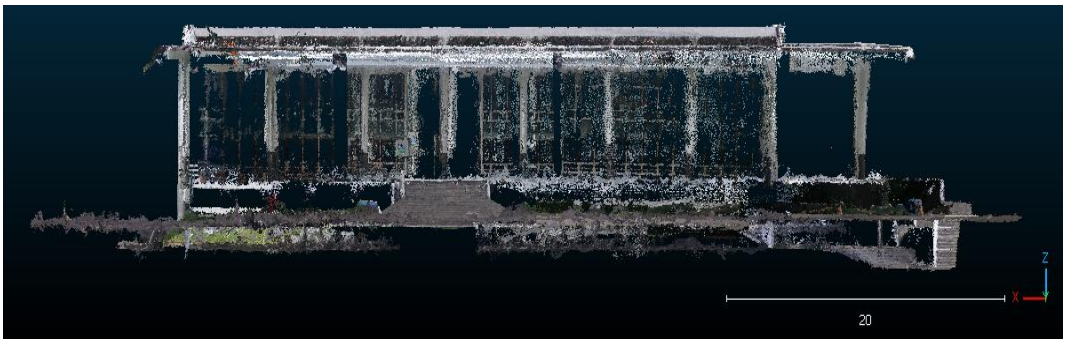

Figure 20. South view point cloud integration result

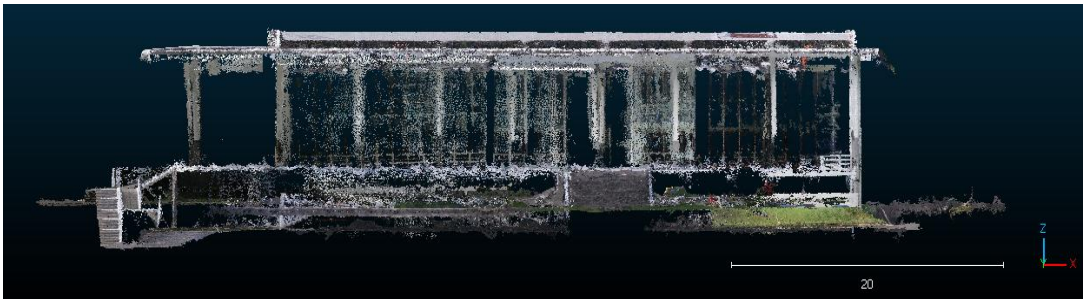

Figure 21. North view point cloud integration result

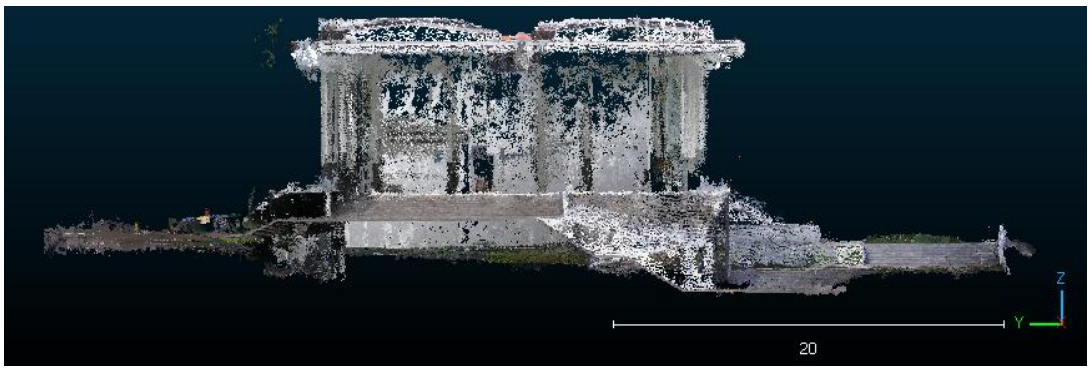

Figure 22. East view point cloud integration result

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

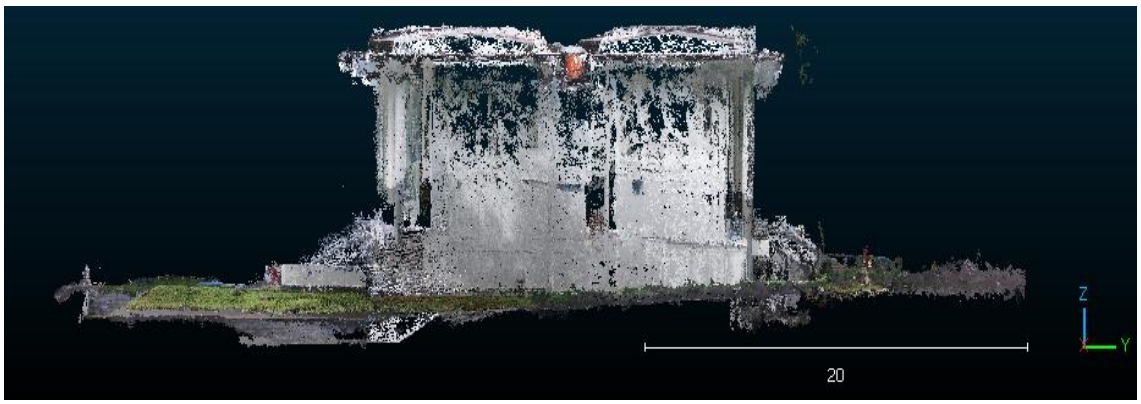

Figure 23. West view point cloud integration result

To evaluate the impact of the integration process, a visual comparison was performed using cloud-to-cloud (C2C) distance analysis between the point clouds before and after integration. In Figure 24., the left image shows the distance values prior to integration, while the right image displays the result after the process.

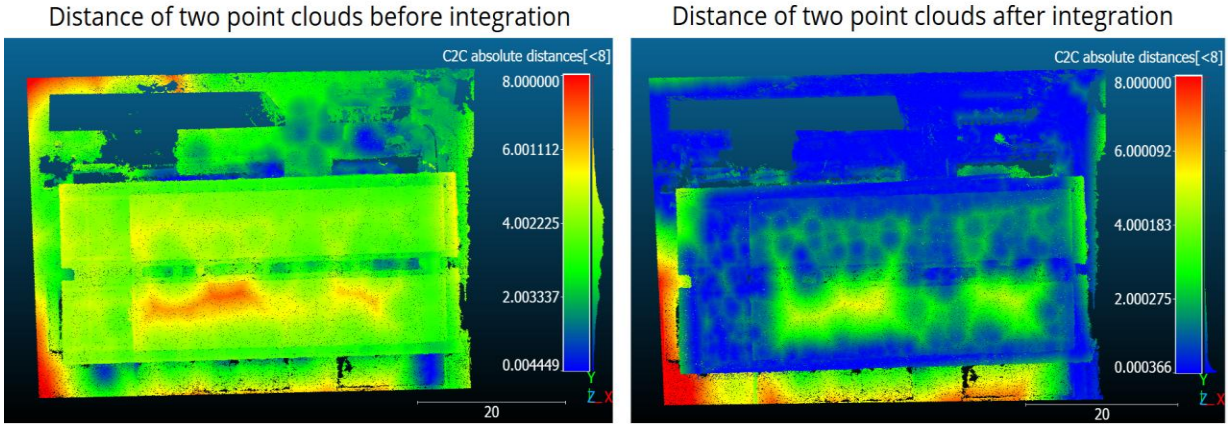

Figure 24. Cloud-to-cloud distance comparison between point clouds before (left) and after (right) integration

The color scale in the visualization represents the spatial distance between the two point clouds. Blue indicates a distance of 0–2 meters, green 2–4 meters, yellow 4–6 meters, and red 6–8 meters. The more blue an area appears, the smaller the difference between point positions in the two datasets. After the integration process, a wider distribution of blue areas is observed, particularly on rooftops and building façades. This indicates improved spatial alignment and greater integration accuracy. The improvement is most evident in areas with overlapping coverage between the UAV imagery point cloud and the action video camera point cloud. In these regions, the same objects were captured from different viewpoints, allowing the integration process to align the point positions more accurately. In contrast, areas that remain yellow to red represent regions without overlap between datasets. Since the same objects were not captured in those zones, the integration had little effect on spatial alignment there.

3.4 Validation

The results of the point cloud integration that have been obtained are then tested for quality through the process of position validation and dimensional validation.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

3.4.1 Position Validation

In the position validation process, a sample of 11 units was used, with the distribution shown in **Figure 25**. The quality of GNSS measurements for each sample is presented in Table 6. where the average HRMS and VRMS for all sample points are 0.015 meters and 0.022 meters, respectively.

Figure 25. Distribution of position validation sample points

Table 6. shows the accuracy of position validation point measurements, where all points have a "Fixed" status, indicating that the position measurements are reliable.

Table 6. Position validation point measurement accuracy

Point Id	HRMS (m)	VRMS (m)	Status
1	0.015	0.027	Fixed
2	0.0117	0.014	Fixed
4	0.0092	0.016	Fixed
5	0.0114	0.017	Fixed
8	0.0122	0.017	Fixed
9	0.012	0.017	Fixed
10	0.0106	0.018	Fixed
11	0.0498	0.072	Fixed
12	0.0086	0.013	Fixed
13	0.0136	0.014	Fixed
14	0.013	0.026	Fixed

The position quality of the point cloud integration results is assessed by the RMSE of the position of each sample, the RMSE of the 3D coordinates (X, Y, and Z), and the total RMSE, which can be seen in **Table 7.** The RMSE values are derived from the points in the point cloud and compared with the results obtained from GNSS. Sample 1 exhibits the largest RMSE value of 1.108 meters. When considering the position of the sample, Sample 1 is located at the edge of the road marking, which is at risk of changing during the point cloud data acquisition and the validation process. This increases the likelihood of errors in point acquisition in both the integration results and during the validation process.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

Table 7. RMSE value of each point

Point Id	X (m)	Y (m)	Z (m)	X′ (m)	Y′(m)	Z′ (m)	RMSE (m)
1	9237506.980	788495.457	788.342	9237506.016	788495.586	788.872	1.108
2	9237484.984	788494.808	789.522	9237484.564	788494.942	789.666	0.463
4	9237501.192	788525.180	788.942	9237500.931	788525.055	788.983	0.292
5	9237505.869	788527.599	788.63	9237505.494	788527.563	788.7	0.384
8	9237504.716	788548.778	788.491	9237504.294	788548.665	788.516	0.437
9	9237471.693	788548.951	788.36	9237471.221	788549.119	788.518	0.525
10	9237469.014	788549.740	788.044	9237468.706	788549.723	788.243	0.367
11	9237475.421	788533.883	790.784	9237475.732	788533.721	790.783	0.350
12	9237468.750	788531.755	788.115	9237468.268	788531.742	788.362	0.542
13	9237470.977	788527.944	788.267	9237470.631	788528.21	788.437	0.469
14	9237470.591	788507.054	789.257	9237470.228	788507.207	789.278	0.394

The quality of the position resulting from the integration of the point cloud, assessed based on the RMSE of each sample, RMSE of the 3D coordinates (X, Y, and Z), and the total RMSE, can be seen in **Table 8.** The RMSE validation position for all samples is 0.528 meters.

Table 8. RMSE of the 3D coordinates (X, Y, and Z) and the total RMSE

$RMSE_X$ (m)	$RMSE_Y$ (m)	$RMSE_Z$ (m)	$RMSE_{Total}$ (m)
0.466	0.139	0.205	0.528

3.4.2 Dimensional Validation

Dimension validation is conducted to ensure the accuracy of the measured object sizes. The dimension validation samples are distributed across the West, North, and South sides, with the main components measured being the supporting column of the building and the width of the stairs. No samples were taken on the East side because the supporting column is not clearly visible. The visualization of the distribution and results of the dimension validation sampling can be seen in **Figure 26.** until **Figure 28.**

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

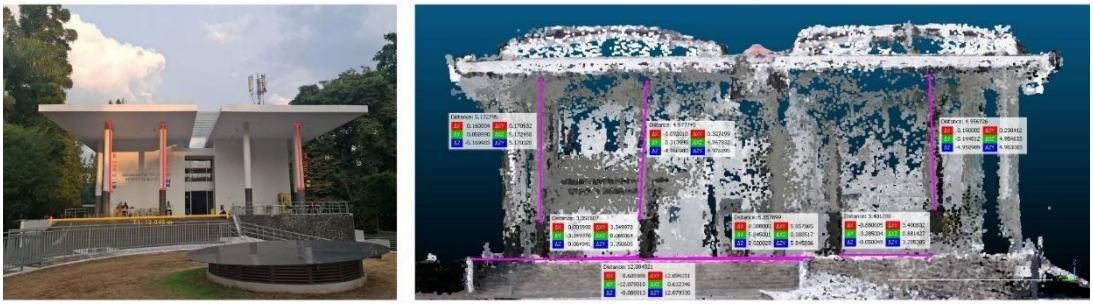

Figure 26. Distribution and results of measurements on the West side dimensional validation sample

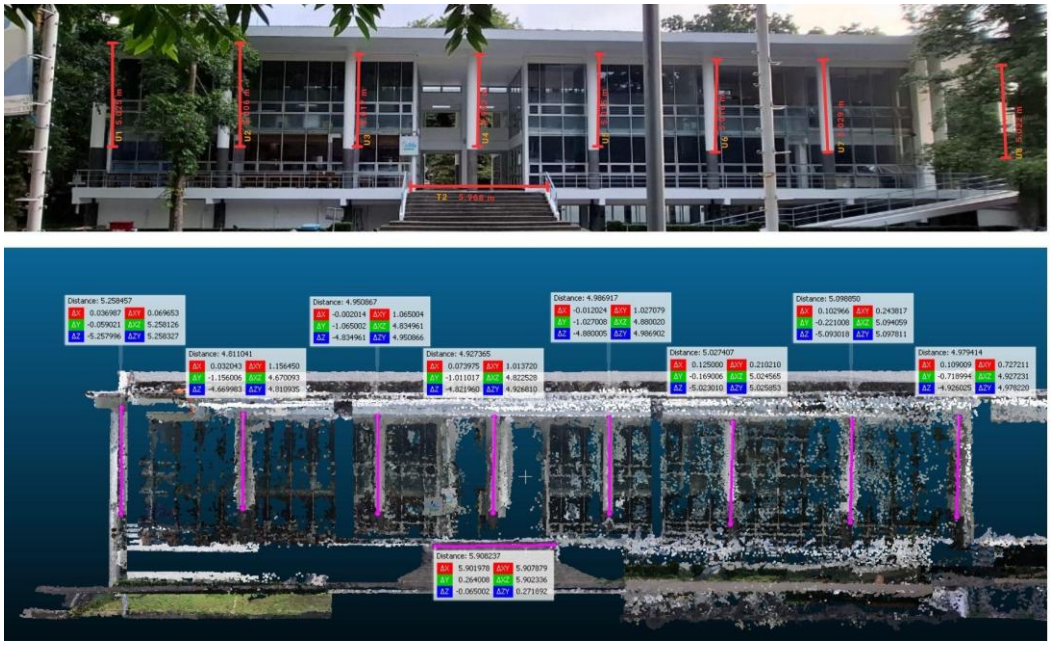

Figure 27. Distribution and measurement results on the North side dimension validation sample

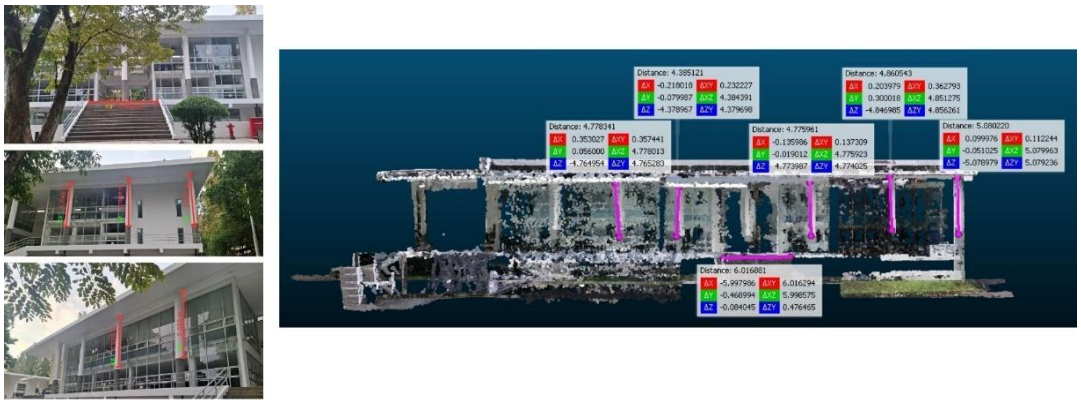

Figure 28. Distribution and measurement results on the South side dimension validation sample

The quality of the dimensions resulting from the integration of the point cloud, assessed by the RMSE of the dimensions between the point cloud and distometer measurements, can be seen in **In the** context of the dimensional validation, dd represents the actual measured dimensions obtained from the reference or true values, dd' denotes the dimensions measured from the point cloud data, and Δ is the difference between the actual dimensions d and the dimensions from the point cloud d', calculated $as \Delta = d - d'$.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

Table 9**.** The RMSE value for the dimensions is 0.578 meters. In the context of the dimensional validation, dd represents the actual measured dimensions obtained from the reference or true values, dd' denotes the dimensions measured from the point cloud data, and Δ is the difference between the actual dimensions d and the dimensions from the point cloud d', calculated as$\Delta = d - d'$.

Table 9. RMSE of dimensional validation

Id	d (m)	d' (m)	Δd (m)
U1	5.025	5.258	0.233
U2	5.006	4.811	0.195
U3	5.017	4.95	0.067
U4	4.993	4.927	0.066
U5	5.015	4.986	0.029
U6	5.016	5.027	0.011
U7	5.029	5.098	0.069
U8	5.022	4.979	0.043
T2	5.968	5.908	0.06
K1	5.002	5.172	0.17
K2	5.026	4.977	0.049
K4	5.01	4.956	0.054
T1	12.045	12.094	0.049
S1	5.025	5.08	0.055
S2	5.026	4.86	0.166
S3	5.04	4.775	0.265
S4	5.035	4.385	0.65
S5	5.035	4.778	0.257
T3	6.046	6.016	0.03
	Σr		**2.518**
	Σr^2		**6.34**
	RMSE		**0.577**

Reviewing the quality of the positional and dimensional results from the integrated point clouds, further utilization such as mapping and 3D modeling can be conducted. The quality of the models produced can achieve Level of Detail (LoD) 2 with a minimum 3D positional accuracy of 1 meter. The requirements for the minimum 3D positional accuracy for each LoD are outlined in the CityGML format standard document by the Open Geospatial Consortium from 2012, which can be found in **Table 10.**

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

Table 10. 3D Modeling standard of the CityGML format (Open Geospatial Consortium, 2012)

Parameter	LoD0	LoD1	LoD2	LoD3
Model Scale	Regional	City/Regional	District	Architectural Model (Exterior)
3D Positional Accuracy	Below LoD1	5 m	2 or 1 m	0.5 m
Object Generalization	Land use classification	> 6*6 m	> 4*4 m	> 2*2 m
Building Parts	-	-	-	External structure
Roof Projection	-	Average	Type and orientation	Actual object shape

4. CONCLUSIONS

The integration of point clouds obtained from aerial photogrammetry and action video cameras provides an effective solution for urban 3D mapping. This study demonstrates that such integration can produce detailed and accurate representations of urban environments, which are crucial for addressing the complexities of city planning and management. The point clouds generated from aerial photogrammetry offer 3D positional accuracy with an RMSE of 0.19073 meters, while the point clouds from the action video camera show RMSE values of approximately 0.487970 meters for geolocation accuracy on the X-axis, 0.393635 meters on the Y-axis, and 0.707503 meters on the Z-axis. The position validation results indicate a total RMSE of 0.52800 meters, and the RMSE value for dimensions is 0.578 meters, confirming that the integration of point clouds from both sources meets the required accuracy standards. With a minimum 3D positional accuracy of 1 meter, this aligns with the Level of Detail (LoD) 2 standards set by the Open Geospatial Consortium 2012. This finding indicates that for this study area, integration yields reliable results; however, in larger and more complex urban environments, the outcomes may differ. Additionally, quality optimization in the filtering process and point cloud integration can be performed to create models with a higher LoD level, specifically LoD 3, which requires a minimum 3D positional accuracy of 0.5 meters. Overall, these findings highlight the importance of integrating photogrammetry and action video technologies to create reliable urban mapping solutions. By combining data from both sources, it is possible to generate more comprehensive and accurate information, greatly supporting decision-making in sustainable urban planning and management.

5. RECOMMENDATIONS

The recommendations from this research are to explore various urban environments, especially in larger and more complex areas, to evaluate the performance of integration methods under different conditions. Additionally, integrate other technological sources such as LiDAR and satellite imagery to enrich the 3D mapping results. Furthermore, conduct a cost-benefit analysis to evaluate the use of this mapping technology compared to existing methods.

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529

6. REFERENCES

Babahajiani, P., Fan, L., Kämäräinen, J. K., and Gabbouj, M. (2017). Urban 3D Segmentation and Modelling from Street View Images and LiDAR Point Clouds. *Machine Vision and Applications*, *28*(7), 679–694. https://doi.org/10.1007/s00138-017-0845-3

Che Ku Abdullah, C. K. A. F., Baharuddin, N. Z. S., Ariff, M. F. M., Majid, Z., Lau, C. L., Yusoff, A. R., Idris, K. M., and Aspuri, A. (2017). Integration of Point Clouds Dataset from Different Sensors. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(2W3), 9–15. https://doi.org/10.5194/isprs-archives-XLII-2-W3-9-2017

Eltner, A., and Sofia, G. (2020). *Structure from motion photogrammetric technique* (pp. 1–24). https://doi.org/10.1016/B978-0-444-64177-9.00001-1

Gonçalves, J. A., and Pinhal, A. (2018). Mobile Mapping System Based on Action Cameras. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(1), 167–171. https://doi.org/10.5194/isprs-archives-XLII-1-167-2018

Gu, X., Wang, X., and Guo, Y. (2020). A Review of Research on udhaPoint Cloud Registration Methods. *IOP Conference Series: Materials Science and Engineering*, *782*(2). https://doi.org/10.1088/1757-899X/782/2/022070

Hapriansyah, S. A. (2021). Analisis Perbandingan Ketelitian Hasil Orthomosaic Menggunakan Perangkat Lunak Komersial Pix4Dmapper dan Open Source WebODM Drone. *Jurnal Teknis ITS*, *10*.

Ihsan, M., and Sugandi, D. (2019). Pemanfaatan Produk Fotogrametri Digital untuk Media Pembelajaran. *Jurnal Geografi Gea*, *19*(2).

Kresna Sumaamijaya, A., and Gilang Attalla, M. (2025). *Study of Quality Control for Mapping Product: Aerial Photogrammetry Data Acquisition Using UAV (Unmanned Aerial Vehicle) Based on SNI 9135-1: 2022*. https://ejournal.upi.edu/index.php/IJGIS/index

Li, D., He, K., Wang, L., and Zhang, D. (2022). Local Feature Extraction Network with High Correspondences for 3d Point Cloud Registration. *Applied Intelligence*, *52*(9), 9638–9649. https://doi.org/10.1007/s10489-021-03055-1

Li, P., Wang, R., Wang, Y., and Tao, W. (2020). Evaluation of the ICP Algorithm in 3D Point Cloud Registration. *IEEE Access*, *8*, 68030–68048. https://doi.org/10.1109/ACCESS.2020.2986470

Nex, F., and Remondino, F. (2014). UAV for 3D mapping applications: a review. *Applied Geomatics*, *6*(1), 1–15. https://doi.org/10.1007/s12518-013-0120-x

Nuryudha, W. (2024). *Analisis Perbandingan Pemodelan 3D Menggunakan Metode Fotogrametri dan Videogrametri Dengan Wahana UAV Quadcopter dan Kamera DSLR (Studi Kasus : Candi Gunung Gangsir, Pasuruan)*.

Pepe, M., Alfio, V. S., Costantino, D., and Herban, S. (2022). Rapid and Accurate Production of 3D Point Cloud via Latest-Generation Sensors in the Field of Cultural Heritage: A Comparison between SLAM and Spherical Videogrammetry. *Heritage*, *5*(3), 1910–1928. https://doi.org/10.3390/heritage5030099

Persad, R. A., and Armenakis, C. (2017). Automatic co-registration of 3D multi-sensor point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 162–186. https://doi.org/10.1016/j.isprsjprs.2017.05.014

Qian-Yi Zhou, Park, J., and Koltun, V. (2016). Fast Global Registration. *Computer Vision–ECCV2016*, *9906*. https://doi.org/10.1007/978-3-319-46475-6

Shahzad, M., and Zhu, X. X. (2015). Robust Reconstruction of Building Facades for Large Areas Using Spaceborne TomoSAR Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(2), 752–769. https://doi.org/10.1109/TGRS.2014.2327391

Sirmacek, B., and Lindenbergh, R. (2014). Accuracy Assessment of Building Point Clouds Automatically Generated from Iphone Images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *40*(5), 547–552. https://doi.org/10.5194/isprsarchives-XL-5-547-2014

Šiško, D. (2022). Application of 3D City Model in Spatial Planning of the City of Zagreb. *FIG Congress*.

Situmorang, P. H. (2019). Calibration of Digital Cameras for Mobile Mapping Purposes. *JGISE: Journal of Geospatial Information Science and Engineering*, *2*(1). https://doi.org/10.22146/jgise.40817

Suziedelyte Visockiene, J., Brucas, D., and Ragauskas, U. (2014). Comparison of UAV Images Processing Softwares. *Journal of Measurement in Engineering*, *2*(2).

Tomljenovic, I., Höfle, B., Tiede, D., and Blaschke, T. (2015). Building Extraction from Airborne Laser Scanning Data: An Analysis of the State of The Art. In *Remote Sensing* (Vol. 7, Issue 4, pp. 3826–3862). MDPI AG. https://doi.org/10.3390/rs70403826

Xiao, W., Cao, H., Tang, M., Zhang, Z., and Chen, N. (2023). 3D urban object change detection from aerial and terrestrial point clouds: A review. *International Journal of Applied Earth Observation and Geoinformation*, *118*, 103258. https://doi.org/10.1016/j.jag.2023.103258

Yang, Z., Wang, X., and Hou, J. (2021). A 4PCS Coarse Registration Algorithm Based on ISS Feature Points. *Chinese Control Conference, CCC*, *2021-July*, 7371–7375. https://doi.org/10.23919/CCC52363.2021.9549486

Zhu, X. X., and Shahzad, M. (2014). Facade Reconstruction Using Multiview Spaceborne TomoSAR Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(6), 3541–3552. https://doi.org/10.1109/TGRS.2013.2273619

DOI: https://doi.org/10.17509/gea.v25i2.81474
p-ISSN 1412-0313 e- ISSN 2549-7529