

AUTOMATIC METHOD FOR GENERATING 3D BUILDING MODELS WITH TEXTURE FROM UAV IMAGES

H.G. Kim¹, W. Yoon¹, S. Rhee¹, T. Kim²

¹ 3DLabs Co. Ltd., 1609-ho, 56 Songdogwahak-ro, Yeonsu-gu, Incheon, Korea – (khanai, wsyoon, ahmkun)@3dlabs.co.kr

² Dept. of Geoinformatic Engineering, Inha University, 100 Inharo, Namgu, Incheon, Korea – tezyd@inha.ac.kr

Commission I, ICWG I/II

KEY WORDS: 3D Building Model, UAV Images, Point Cloud, Statistical Technique.

ABSTRACT:

In this paper, we propose a method for automatically generating 3D building models from UAV image based point cloud data and for mapping building texture from UAV images. The proposed method generates dense point clouds from UAV images and isolates points from building areas through a statistical analysis. It then generates solid models as 3D building models by point cloud analysis per building area. Texture for 3D building solid models are created by mapping model face to UAV images. In order to verify the proposed method, various UAV image sets and point clouds were tested. As results, the possibility of generating cluster-type solid building models based on UAV images was confirmed. It is expected that this method can contribute to the active usage of UAV images in 3D spatial information generation. In the future, we plan to conduct research on improving the accuracy of curved building shapes and texturing accuracy.

1. INTRODUCTION

1.1 Background

With the recent development of information and communication technology and high precision 2D and 3D data acquisition technology, technologies to build photo-realistic 3D city models for smart cities, digital twins and metaverses, have become highly in demand.

In general, 3D building models can be created by supplementing heights onto 2D map data or by processing 3D data such as a point cloud. Building models are created from 2D data such as a numerical map or an open street map by assigning a height value to the boundary information of a building. Such models can be used as a base data for public institutions or local governments (Ham, 2019). However, there is a limit to modelling of a new building due to long map generation and update cycle, and the need for additional height data or user intervention (Kim *et al.*, 2019).

Building modelling by processing 3D data such as point clouds are also being studied in various fields. Constructing a mesh by applying interpolation technology to a point cloud was conducted. This technique is one of the modelling techniques being actively studied to the extent that it is most used in commercial software (Wu *et al.*, 2017). The mesh model generation technology has the advantage of being able to create a model that best reflects the shape of a point cloud. On the other hand, it is difficult to create a model for each object because building and non-building data are mixed in a point cloud. It is difficult to create a model according to LOD (Level of Detail). a solid model by manually extracting the vertices of buildings from 3D data (Wu *et al.*, 2017) or automatically extracting the intersections of planes and planes through planarity analysis (Nan and Wonka, 2017) were studied. The above methods create a building using the minimum floor plan. Therefore, there are advantages in terms of data management and model update. However, it may be difficult to apply to data containing a relatively large amount of noise because the research was conducted mainly manually-derived or LiDAR-based data with little or no noise. Also, as with the mesh model,

a prerequisite was required to separate building and non-building point clouds. In order to solve this limitation, research on noise removal and sharpening of point clouds was carried out (Sun *et al.*, 2020). In addition, technologies for generating 3D models or point clouds using images are being developed in deep learning fields such as Neural Radiance Fields (NERF) for building modelling (Pumarola *et al.*, 2021).

Texturing building models is one of the crucial factors for the quality of building models. Oh *et al.* (2007) reported a research on how to perform texture mapping and create a realistic 3D building model by using spatial images such as aerial or satellite images on a 3D city model of a wide area quickly and economically. Bulatov *et al.* (2014) reported a research of giving textures to a building model through sensor posture estimation of aerial photos.

In addition, point cloud generation and building model generation technologies are being actively released as software for use. For example ContextCapture by Bentley, SURE by nFrames and Pix4DMapper by Pix4D, etc (Wahbeh *et al.*, 2021). All of these software demonstrate the technology to create high-quality point clouds and building models with outstanding performance. However, there are limitations in that it creates a model that cannot separate the building and the ground, and that some SWs are products developed exclusively for a specific platform containing Lidar data.

While above mentioned research focused on using LIDAR-driven point clouds and aerial photos, we have been focusing on research of using UAV (Unmanned Aerial Vehicle) images for creating point clouds and textured building models. Lim *et al.*, (2019) reported the effects of UAV image quality on the performance of point cloud generation. Stereo-plotting and mapping accuracy of the UAV was recently performed (Kim *et al.*, 2018; Lee *et al.*, 2018) for application in the field of spatial information. In this paper, we report technology developed for automated generation of 3D building models from UAV-driven point clouds and for high quality building texture mapping from UAV images. To evaluate the applicability of the developed method, we report 3D building models automatically generated from various UAV image-based point clouds.

2. METHOD

2.1 Proposed Method

In this paper, we propose automatic method for generating clustered 3D building models from UAV image based point clouds. The process of the proposed method is shown in the following Figure 1.

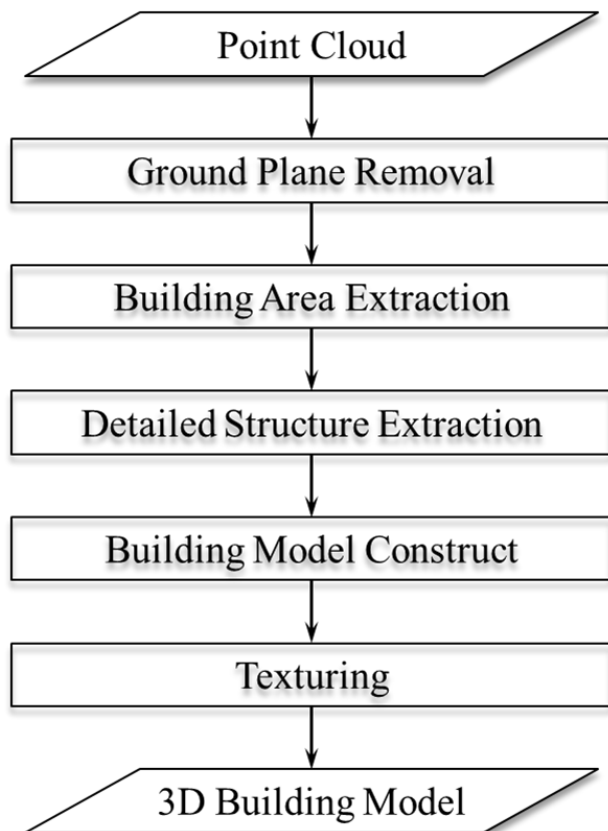


Figure 1. Process of proposed method.

The proposed method consists of five steps. First, the input point cloud is analysed by a statistical method and points corresponding to the ground are removed. In the second step,

the building area is extracted using the point cloud from which the ground has been removed. In the third step, the detailed structure area existing within the building area is extracted. In the fourth step, using the building, structure, and ground height, the outer point of the building is constructed and the shape of the 3D building model is created. In the final step, a 3D solid building model is created by extracting the textures corresponding to each side of the building model shape from the UAV images.

2.2 Ground Plane Removal

In order to generate clustered 3D building models from point cloud, point separation between the ground and non-ground must be preceded. For ground plane removal, a research was conducted to explore the ground plane through point cloud planarity analysis using a RANSAC (Random Sample Consensus)-based model fitting algorithm (Fischler and Bolles, 1981). In this study, the ground plane is calculated in the point cloud using the RANSAC-based planarity analysis method and the ground and non-ground points are separated. Figure 2 shows the ground plane removal process within the point cloud and Figure 3 shows the result.

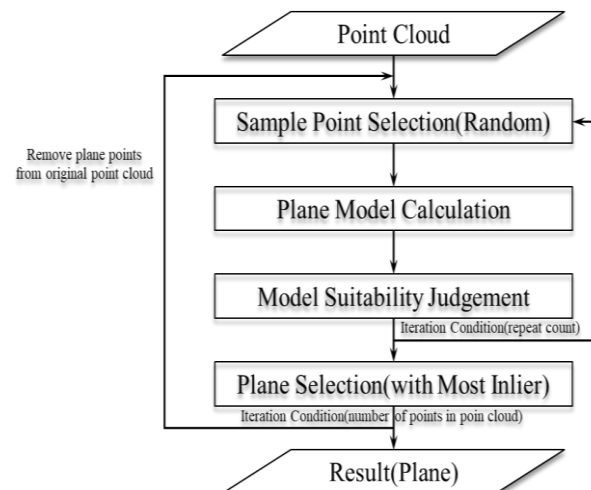


Figure 2. Process of ground plane removal.

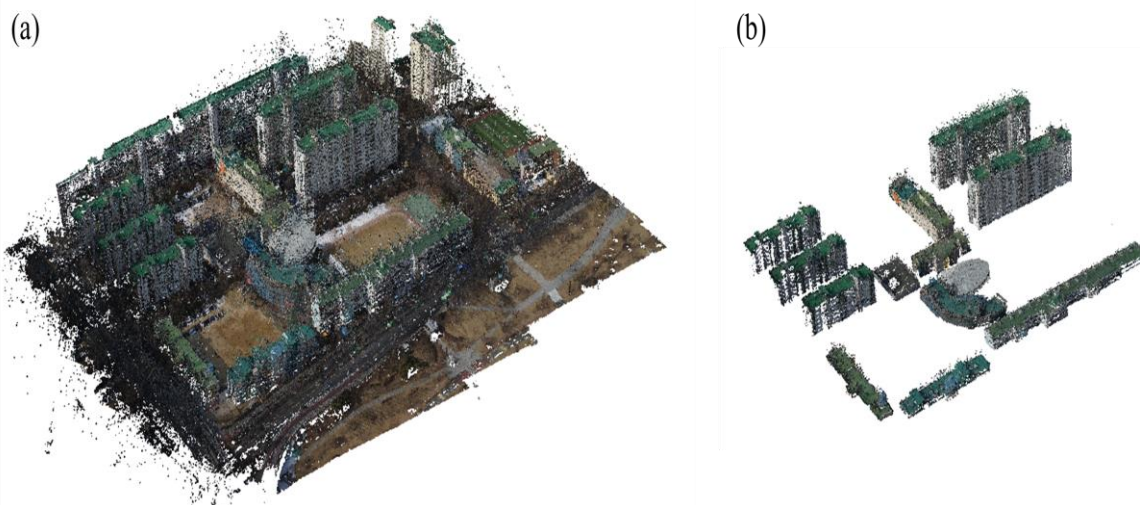


Figure 3. Source point cloud(a) and Ground plane removal result(b).

2.3 Building Area Extraction

In the second stage, a research was performed to extract the building area by converting the point cloud from which the ground was removed to an ortho image (Kim *et al.*, 2019). The process is shown in the following Figure 4.

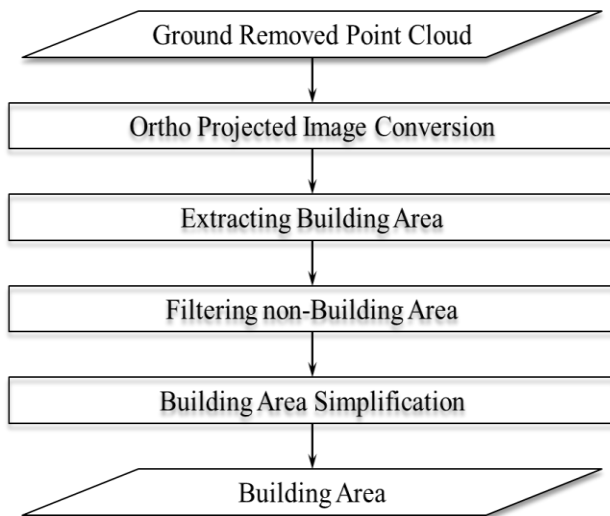


Figure 4. Process of building area extraction.

This method consists of converting non-ground point cloud with some noises into an ortho image and extraction of the building area from the image using the image processing. The process is shown in the following Figure 5.

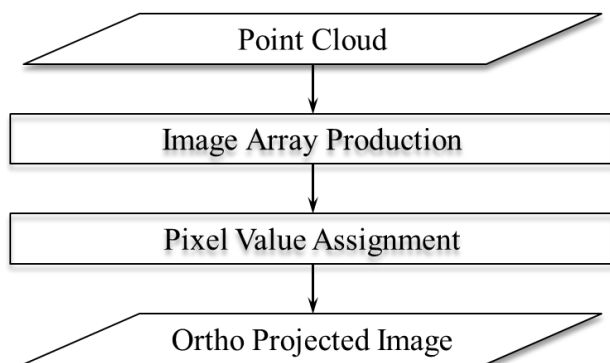


Figure 5. Process of converting into an ortho image

The first step, ortho image conversion is performed in the following way. First, an area is set using the maximum and minimum values of the ground coordinates x and y of the point cloud. Then, the set area is divided into grid units of a certain GSD (Ground Sample Distance) size using a point cloud. After that, the corresponding points are selected for the ground coordinates (X , Y) of all generated grids, and the color (Red, Green, Blue) or height value (Z) of the selected point is applied. In this case, the points corresponding to each grid are calculated using a sensor model equation such as a collinear condition equation or DLT (Direct Linear Transformation). When converted to an ortho image in this way, a building area having three-dimensional coordinates can be extracted using image processing, and it has an advantage over handling a point cloud directly in terms of processing speed. In addition, by designating the brightness value of the ortho image as the Z value, there is an advantage that the height value can be easily specified for each pixel in the building area.

The image processing techniques used in this paper include techniques to fill in noise and remove holes in the ortho image, such as dilation and eroding, and contouring to extract the building area. The above techniques were performed using functions such as `dilate`, `erode`, and `findContours` of `opencv`. In this case, the height value of the point is used as the value of the ortho image. There are two reasons for extracting the building area by converting the point cloud into an image. The first is that converting 3D data into 2D data enables easy and fast data processing, and the second is to facilitate the extraction of height values for each building. In this study, the building area was extracted using the research.

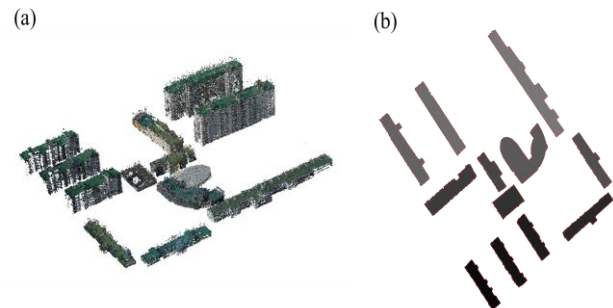


Figure 6. Ground removed point cloud(a) and initial building area in ortho image.

The building area created by this method consists of all points surrounding the building area. If we construct the shape of a building with many points, one face may not be composed of one polygon, but several polygons. In this case, it increases the processing time and causes distortion in the texturing step. Therefore, it is necessary to simplify the area to the minimum number of points that represents the building area. Figure 7 shows the result of simplification.

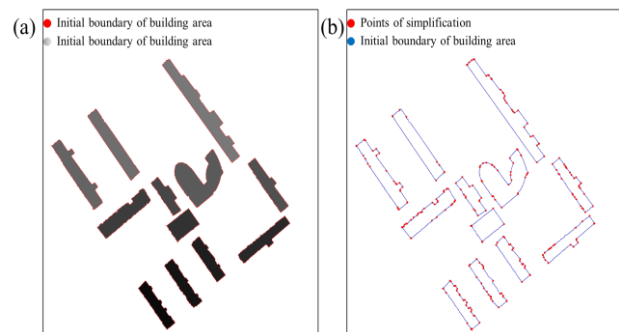


Figure 7. Building area(a) and simplification result(b).

The building area point simplification is performed through line fitting algorithm based on RANSAC. First, the line composed of the building area is searched. After that, the line with the largest number of inlier points is sequentially selected as the optimal building area.

2.4 Detailed Structure Extraction

Extracting the detailed structure of the building is the step of composing the internal structures within the body of the building. First, the height of the roof is determined by analysing the value for each building area on the previously generated ortho image. Then, the image is binarized based on that height, and the area corresponding to the detailed area higher than the roof is separated. The separated detailed structure area is

extracted as an optimal detailed structure area through the process of area extraction and simplification similar to the building area extraction method.

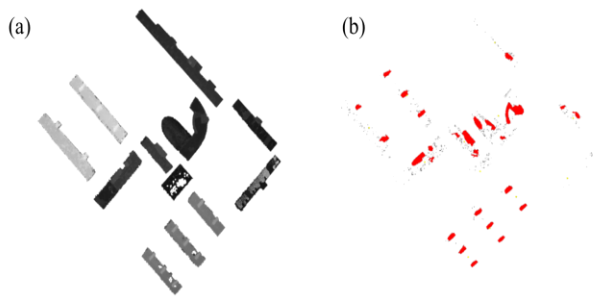


Figure 8. Ortho image(a) and detailed structure area(b).

2.5 Building Model Construct

In this step, the shape of the building model is created by substituting 3D coordinates X, Y, and Z for the points of the building and detailed structures. The 3D coordinates of the building are calculated using column and row information of the building area and detailed structure extracted in the previous step. The 2D plane coordinates X and Y are calculated by transforming the columns and rows using the formula for converting the point cloud to the ortho image. And we use the mode of brightness value of the area in the ortho image as the height value Z. Each point uses the actual height value at that point, so it is possible to model even a sloping building. At this time, we use the ground plane height for the ground height value around a building, and the roof height value of the building area for the floor height of the building detail structure. Through the above processes, the LOD (Levels of Details) of the building model targeted in this paper is LOD2.2 based on CityGML.

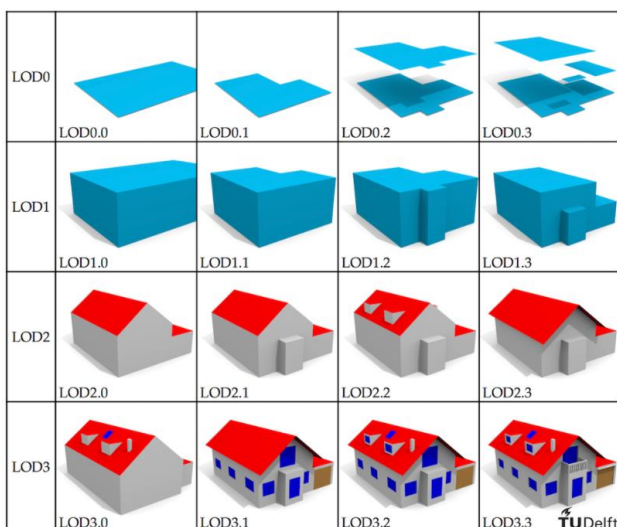


Figure 9. Targeted levels of detail (Biljecki *et al.*, 2016).

2.6 Texturing

Finally, textures are applied to the shape of the building model structure to generate a complete textured 3D building model. Texturing is performed by extracting the texture of each side within the image area using the image and sensor-modelled orientation parameters for each face of the previously created building shape. The process is shown in the following Figure 10.

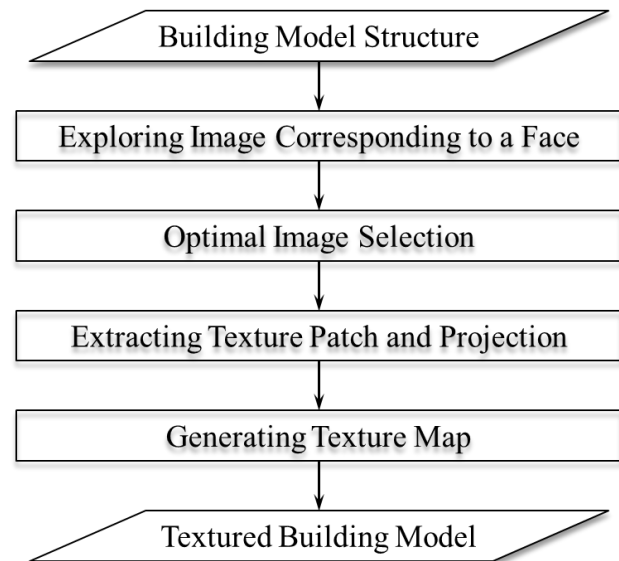


Figure 10. Process of texturing.

First, images of each face are found by using the coordinates of the four vertices of the building face and orientation parameters of each image. Second, an optimal image that satisfies the area and angle conditions of the projected surface from among the images in which each surface is captured is selected. For the optimal image, the image with the largest area within the image and the angle calculated using the normal vector is closest to the vertical direction to the looking angle is selected. Third, a patch extracted in the selected image of face is projected and distortion is corrected by applying projection transformation. By applying the above method to all building areas extracted within the point cloud, clustered solid building models can be generated. The Figure 12 shows that the results of texturing.

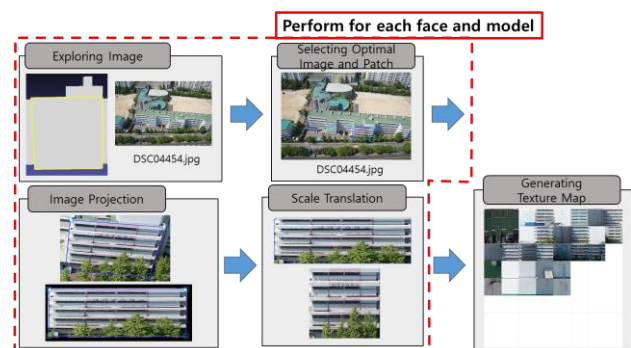


Figure 11. Process of texturing by pictures.

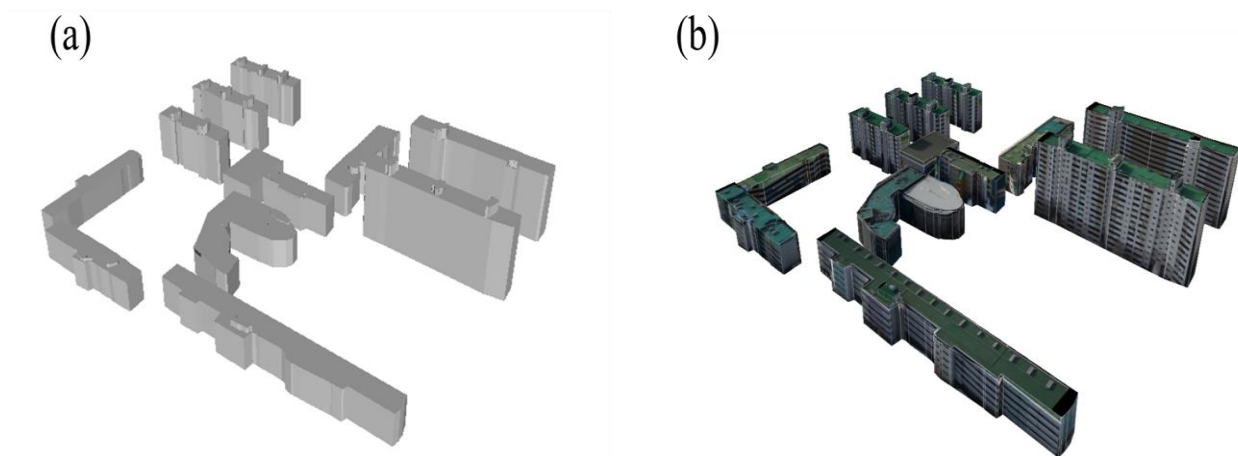


Figure 12. Building structure(a) and texturing results(b).

3. EXPERIMENTS

To evaluate the method proposed in this paper, the proposed method was applied to various types of point clouds and the results were confirmed.

3.1 Experiment Data

In the experiment, we used several UAV images and point clouds generated from them. We used UAV images of Seoul, Siheung, and Goyang, Korea by using a UAV with a Sony ILCE-6000 camera and Incheon by using DJI's Inspire2 using Zenmuse X5S camera. We generated point clouds by using an in-house SW, 3D-UAV, and a commercial SW, Pix4D Mapper. The point cloud was extracted to have a point density at intervals of GSD x 7-8 and no editing was performed. That is, our all UAV images have a GSD of 2-3cm, and the point cloud have a density containing 100 - 400 points per m². Figure 13 show point clouds extracted for the experiments. The data capacity for each experimental data is as follows Table1.

	Seoul	Siheung1	Siheung2	Goyang	Incheon
Data Capacity	1,039,284	117,227	553,533	181,159	320,458

Table 1. Experiment data capacity (KB)

3.2 Evaluation Method

The evaluation of the proposed method was performed by checking whether all buildings of the size larger than the extraction target existing on the point cloud were created. And how much error each vertex of the model had with the vertex coordinates of the point cloud was measured. The reason for not measuring the error with GCP is that the model generated through this study shows the point cloud dependent accuracy. In addition, the processing time for each step was measured. The experiment was conducted in a desktop pc environment including Intel Corei7-9, 9700 CPU, GeForce RTX 2080 D6 8GB GPU and DDR4 32G RAM.

3.3 Experiment Results

In this study, the minimum height of the building to be extracted was set to 6m. Assuming that the height of one floor is 3 m, this was to extract buildings with two or more floors. In addition, the width of the building was extracted for buildings of 100 square meters or more. The figure 14 shows that the experiments results.

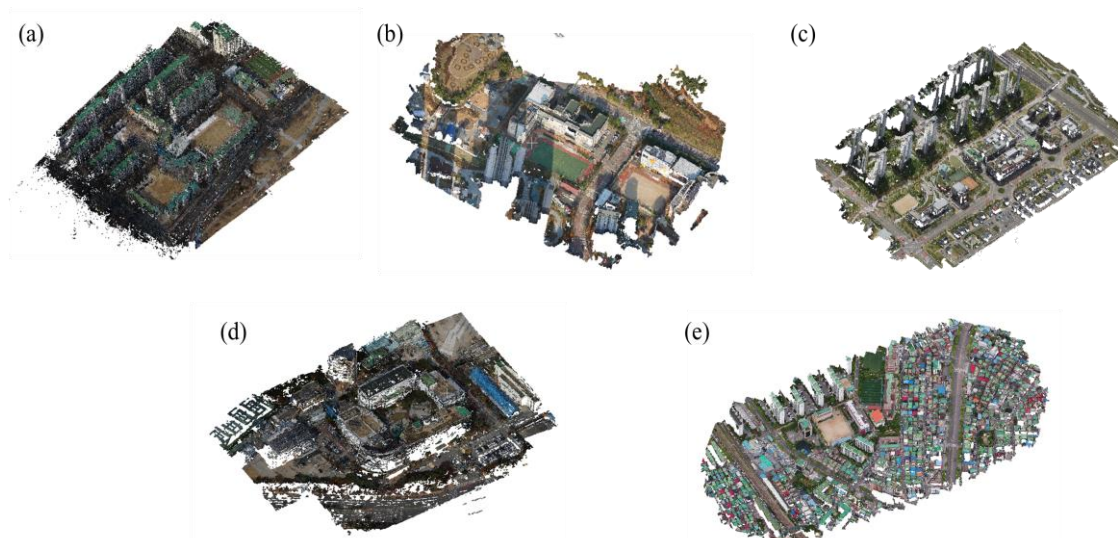


Figure 13. Experiments data. (a)Seoul, (b)Siheung1, (c)Siheung2, (d)Goyang, (e)Incheon.

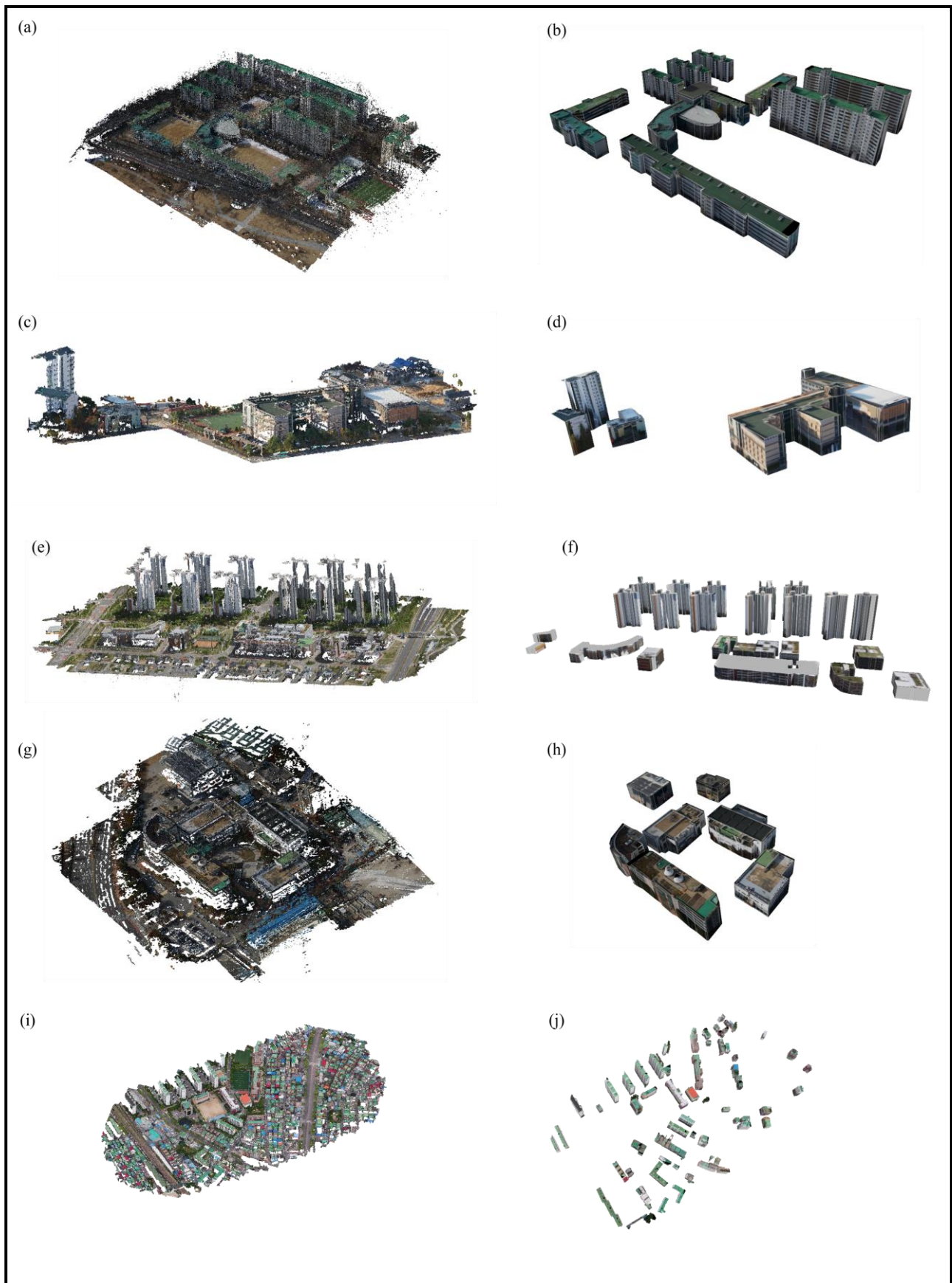


Figure 14. Building Generation Results. (a)Seoul point cloud, (b)Seoul building models, (c) Siheung1 point cloud, (d) Siheung1 building models, (e) Siheung2 point cloud, (f) Siheung2 building models, (g) Goyang point cloud, (h) Goyang building models, (i) Incheon point cloud, (j) Incheon building models.

Through the experiment, it was confirmed that models for all buildings targeted from each data were generated. However, when houses were clustered together at narrow intervals, such as in Incheon, an error occurred in which several buildings were created as one model. This point is a current limitation of this study, and when the ground is removed, the same phenomenon occurs when the two buildings were extracted in the form of attached ones. Also, when buildings did not satisfy the size constraints like Incheon data, they were not modelled.

Finally, although the LOD of the building model pursued in this paper is 2.2, it was not confirmed whether the oblique building model was modeled due to the limitation of the dataset. The following is an accuracy evaluation using the difference between the vertex coordinates of the model and the point cloud. The figure 15 below shows the coordinate measurement method, and the table shows the RMSE(Root Mean Square Error) for each data.

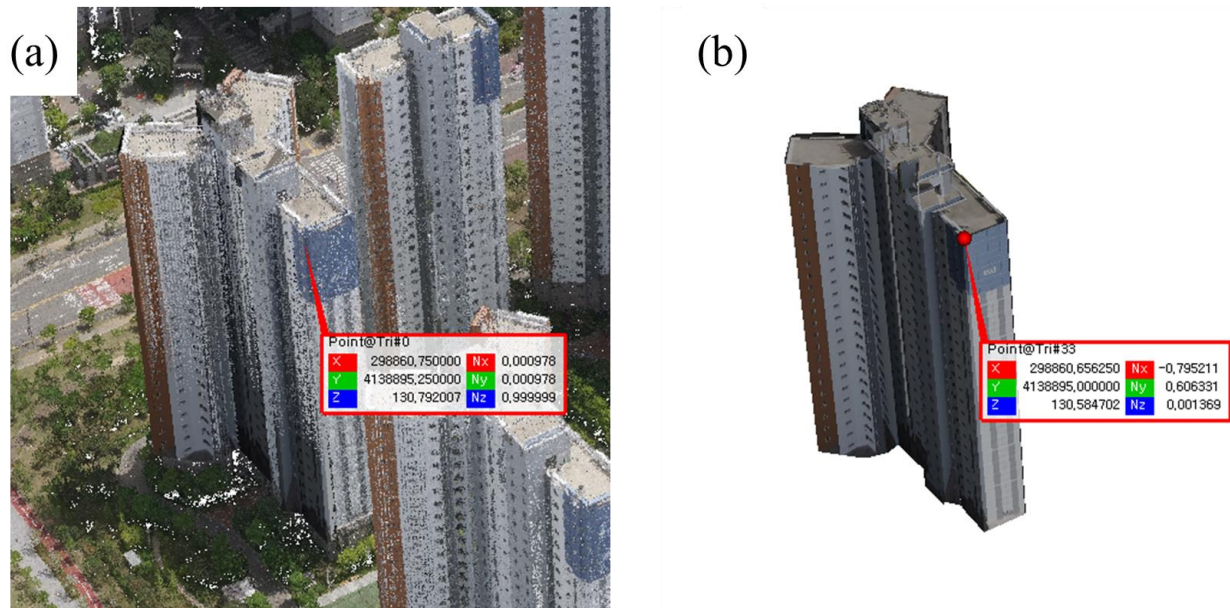


Figure 15. Accuracy Measurement. (a) Point cloud, (b) Model

	Seoul	Siheung1	Siheung2	Goyang	Incheon
Building01	1.22	1.27	0.31	1.01	0.4
Building02	1.15	1.33	0.2	0.63	0.42
Building03	0.94	1.25	0.34	0.85	0.31
Building04	0.87	1.24	0.29	0.78	0.47
Building05	0.94	-	0.32	1.01	0.64
Building06	1.11	-	0.38	0.83	0.42
Building07	1.03	-	0.34	0.26	0.49
Building08	1.34	-	0.28	-	0.87
Building09	0.89	-	0.29	-	0.46
Building10	0.84	-	0.24	-	0.21
⋮	⋮	-	⋮	-	⋮
RMSE(m)	1.02	1.27	0.30	0.81	0.44

Table 2. Accuracy for each dataset

From the above result, it can be confirmed that the error between the point cloud and the building model is determined to be around 1m. It is estimated that the error occurs during the image conversion and image processing of the point cloud. In particular, it is thought to be a phenomenon that occurs when the actual location of the point is interpolated to the outside while applying the dilation and erosion techniques to the ortho image.

The processing time for each data used in the experiment is shown in the table below.

	Seoul	Siheung1	Siheung2	Goyang	Incheon
Ground Plane Removal	577.0	6.5	57.2	97.4	53.1
Building Area Extraction	133.8	9.8	140.0	15.0	32.4
Detailed Structure Extraction	5.1	1.2	15.8	1.3	4.5
Building Model Construction	0.1	0.1	0.1	0.1	0.1
Texturing	1,757.3	305.8	981.9	186.0	268.0
Total	2,473.3	323.4	1,195.0	299.8	358.1

Table 3. Processing time (in second).

As seen from the above results, it can be confirmed that the processing time increases in proportion to the experimental data capacity, and that the texturing and ground removal processes occupy most of the processing time. In particular, texturing took up a minimum of 70% and a maximum of 95% of the processing time overall.

4. CONCLUSIONS

As a result of performing automatic building modeling on various point cloud data, it was confirmed that all buildings existing in the data were extracted. It is expected that our method enables shortened processing time and automated processing procedure compared to the existing 3D building modeling technology. The technology developed through this paper can be used to the construction of 3D city models. However, in the experimental data used in this paper, experiments on various shapes such as curved and inclined building models were not carried out. In the future, we plan to develop and experiment technologies that can create models of various types, such as curved and complex models, and can extract precise model shapes and textures.

ACKNOWLEDGEMENTS

This research was supported by a grant (22DRMS-C147287-05) for the development of a customized realistic 3D geospatial information update and utilization technology based on consumer demand, funded by the Ministry of Land, Infrastructure and Transport of the Korean government.

REFERENCES

- Bulatov, D., Häufel, G., Meidow, J., Pohl, M., Solbrig, P. and Wernerus, P., 2014. Contextbased automatic reconstruction and texturing of 3D urban terrain for quick-response tasks, *ISPRS Journal of Photogrammetry and Remote Sensing*, 157-170.
- Biljecki, F., Ledoux, H. and Stoter, J., 2016. An improved LOD specification for 3D building models, *Computers, Environment, and Urban Systems*, 59, 25-37.
- Fischler, M.A. and Bolles, R.C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communication of the ACM*, 4, 381–395.
- Ham, S., 2019. A Study for Providing 3D Building Information based on National Spatial Information, *Conference of Architectural Institute of Korea*, 39, 136.
- Kim, H.G., Hwang, Y. and Rhee, S.A., 2019. Automatic Building Modeling Method Using Planar Analysis of Point Clouds from Unmanned Aerial Vehicles, *The Korean Journal of Remote Sensing*, 35(6), 973-985.
- Kim, T., Lim, P. C., Son, J. and Seo, J. 2018. Feasibility Study of 1:1,000 Scale Map Generation using various UAVs and processing SW, *Proc. of 2018 Korean Society for Geospatial Information System, Busan, Korea*, May 31 - Jun. 1, 15-16.
- Lim, P.C., Kim, H.G. Park J. and Rhee, S.A., 2020. Evaluation of Possibility of Large-scale Digital Map through Precision
- Nan, L. and Wonka, P., 2017. Polygonal Surface Reconstruction From Point Clouds, *The IEEE International Conference on Computer Vision (ICCV)*, 2353-2361.
- Oh, J.H., Shin, S.W., Park, J.H., and Lee, H.S., 2007. 3D Building Model Texture Extraction from Multiple Spatial Imagery for 3D City Modeling, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, 25(4), 347-354.
- Pumarola, A., Corona, E., Pons-Moll, G. and Moreno-Noguer, F., 2021. D-nerf: Neural radiance fields for dynamic scenes, *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10318-10327.
- Sensor Modeling of UAV, *The Korean Journal of Remote Sensing*, 36(6), 1393-1405.
- Sun, Y., Wang, Y., Liu, Z., Siegel, J.E. and Sarma, S.E., 2020, PointGrow: Autoregressively Learned Point Cloud Generation with Self-Attention, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 61-70.
- Wahbeh, W., Ammann, M., Nebiker, S., van Eggermond, M. and Erath, A., 2021. Image-Based Reality-Capturing and 3d Modelling for the Creation of VR Cycling Simulations, *ISPRS Annals of the Photogrammetry*, 4, 225-232.
- Wu, B., Yu, B., Wu, Q., Yao, S., Zhao, F., Mao, W. and Wu, J., 2017. A Graph-Based Approach for 3D Building Model Reconstruction from Airborne LiDAR Point Clouds, *Remote Sensing*, 9(1)