AUTOMATIC GENERATION OF BUILDING MODELS IN DENSE URBAN AREAS USING AIRBORNE LIDAR AND AERIAL PHOTOGRAPH

Junichi SUSAKI

Associate Professor, Department of Civil and Earth Resource Engineering,
Graduate School of Engineering, Kyoto University

Kyotodaigaku-Katsura, Nishikyo-ku, Kyoto, 615-8540, Japan;
Tel&Fax: 81-75-383-3300, E-mail: susaki.junichi.3r@kyoto-u.ac.jp

KEY WORDS: Modeling, LiDAR, aerial photograph, segmentation

Abstract: In this paper, an algorithm is proposed for automatically generating three-dimensional (3D) building models in dense urban areas. Automatic 3D building modeling in dense urban areas is challenging because, especially in Japan, houses that have slant roofs are located close to each other, and their heights are similar. For this case, difficulty in separating point clouds into individual buildings is an obstacle to modeling. To resolve this issue, the proposed algorithm uses the results of building segmentation from aerial photographs. Segmentation of buildings in urban areas, especially dense urban areas, by using remotely sensed images is also challenging because of the unclear boundaries between buildings and the shadows cast by neighboring buildings. The proposed algorithm successfully segments buildings from aerial photographs, including shadowed buildings in dense urban areas. The main factors in successful segmentation of shadowed roofs are (1) combination of different quantization results, (2) selection of buildings according to the rectangular index, and (3) edge completion by the inclusion of non-edge pixels that have a high probability of being edges. On the other hand, filtered airborne light detection and ranging (LiDAR) data are classified into small groups. By considering the segmented regions and the normals, models of actual building types—gable-roof, hip-roof, flat-roof and slant-roof buildings—are generated. To study the accuracy of the modeling, the proposed algorithm is applied to areas of Higashiyama ward, Kyoto, Japan. Owing to the information of building regions provided by segmentation, the modeling is successful even in dense urban areas. Therefore, the proposed algorithm is concluded to be effective in automatically generating building models in dense urban areas.

1. INTRODUCTION

Building models are required in various fields such as urban planning, landscape assessment, and disaster prevention and mitigation. Airborne light detection and ranging (LiDAR) systems are capable of providing the three-dimensional (3D) data necessary for generating the models. However, to the author's knowledge, automatic 3D building modeling in dense urban areas, which is the focus of the author′s research, has been much less reported, while automatic 3D modeling of isolated individual buildings is quite popular (Sampath and Shan, 2010) (Kim and Shan, 2011). For example, in a dense urban area in Kyoto, Japan, traditional Japanese houses are located close to each other, and the heights of the ridges on the roofs are similar. Even if they are located close to one another, modeling of flat buildings in dense urban areas might be possible by current techniques, since the heights of the building are often not similar and differences in height can be easily detected. In contrast, houses having similar heights might be difficult to model using existing algorithms because of the extra complexity in separating point clouds into individual houses.

On the other hand, fusion of airborne LiDAR data with aerial images or satellite images (Sohn and Dowman, 2007) (Rottensteiner, *et al.*, 2007) (Awrangjeb, *et al.*, 2010) has been examined to implement effective modeling. Object-based segmentation is a conventional approach to extracting building geolocation and shape information from images. This segmentation technique is widely used in a number of applications and its use for urban areas has been reported (Lecumberry, *et al.*, 2010) (Tzotsos, *et al.*, 2011). However, object-based segmentation, obtained by using commercial software (ENVI EX Version 4.8), produced poor results when tested on a 25 cm spatial resolution image of the study area. Factors in the unsuccessful segmentation were (1) unclear building boundaries caused by roofs with similar brightness being located close to each other, (2) shadows cast by neighboring buildings and (3) roofs having a rough texture. By considering these factors, an algorithm has been proposed by the author that segments buildings, including shadowed buildings, in dense urban areas from aerial images (Susaki, 2012a).

In the current paper, another algorithm is proposed for automatic generation of 3D building models for dense urban areas, by using the results of the building segmentation from aerial images produced by the aforementioned algorithm. In addition to information about the building area, normals to the roofs are determined on a group of

buildings basis to reflect a spatially linear arrangement. By taking account of the segmented regions and the normals, the correct building types are selected and the models are accurately generated. The locations used in the study, and the data collected from these sites, are described in Section 2. An overview of the previously proposed segmentation algorithm is given in Section 3, along with an explanation of the newly proposed method. The experimental results are reported in Section 4. The results and the validity of the algorithm are then discussed in Section 5. Finally, Section 6 concludes the paper.

2. METHODS

The proposed modeling framework consists of two major processing: segmentation of regions from aerial photograph and building modeling using the segmented results and the filtered airborne LiDAR data.

(1) Segmentation of regions from aerial photograph

The segmentation algorithm used in the current work was developed by the author (Susaki, 2012a). The flowchart of the algorithm is shown in Figure 1 and the summary of the algorithm is given.

The algorithm can segment buildings, including shadowed buildings, in dense urban areas from an aerial image. To deal with roofs having a rough texture, digital number (DN) is quantized into several ranges of values. A few widths of quantization are applied, and homogeneous areas are labeled. Then, edges extracted from homogeneous areas at each quantization width are merged, and frequently observed edges are extracted. Using a "rectangular index", shown in Figure 2, regions whose shapes are close to rectangular are selected.

The algorithm merges regions segmented using the edges detected with different DN interval widths. This quantization is a type of spatial filtering, and the process is similar to that of smoothing with different spatial scales and merging the results. However, unlike traditional popular smoothing filters, here the edges are preserved and, importantly, the scale of spatial filtering is optimized with respect to the size of building roofs in a locality.

The experimental results from different types of dense urban areas show that the proposed algorithm generated more reasonable and useful segmentation results than did an existing algorithm such as ENVI EX. The main factors in successful segmentation of shadowed roofs are (1) the combination of the results at different quantizations, (2) selection according to the rectangular index and (3) edge completion by the inclusion of non-edge pixels that have a high probability of being edges. With these factors, the proposed algorithm optimizes the spatial filtering scale to the size of local building roofs.

(2) Building modeling

The framework of the building modeling algorithm is shown in Figure 3. While segmented results can help in modeling, the results may be imperfect. Some roofs may be missing, some roofs belonging to the same building may be unified, or some roofs belonging to neighboring buildings may be falsely segmented. Hereafter, the technique to overcome such improper segmented results is introduced.

The building regions are classified into different normal types based on the normals calculated in Step (1). Left figure of Figure 4 shows that both Regions α and β have a single normal direction, whereas Region γ has two major normals. For each region, a probability distribution of the five normal directions of the points (vertical and the four cardinal directions) included in the region is generated. If the maximum probability is larger than a given threshold, the region is assigned a single normal class and the regional "rough direction" is determined. The exact normal is calculated by selecting the normals that are closest to the regional rough direction. If the maximum probability is less than the threshold and the sum of the two highest probabilities is larger than another threshold, the region is given a "mixed normal" classification.

Then, the algorithm searches for a "normal pair" partner for a region that has a single normal and that has not been assigned a "vertical class." The search is along the direction opposite to the azimuthal direction of the regional normal. If another region is detected that exceeds a threshold value, the search is terminated. The search is also terminated if no region is detected that exceeds another threshold. The partner is determined to have been found when the detected region has a single normal class and the normal direction is linearly symmetric to the original normal direction, or when the detected region has a mixed normal class. However, the ratio between the maximum

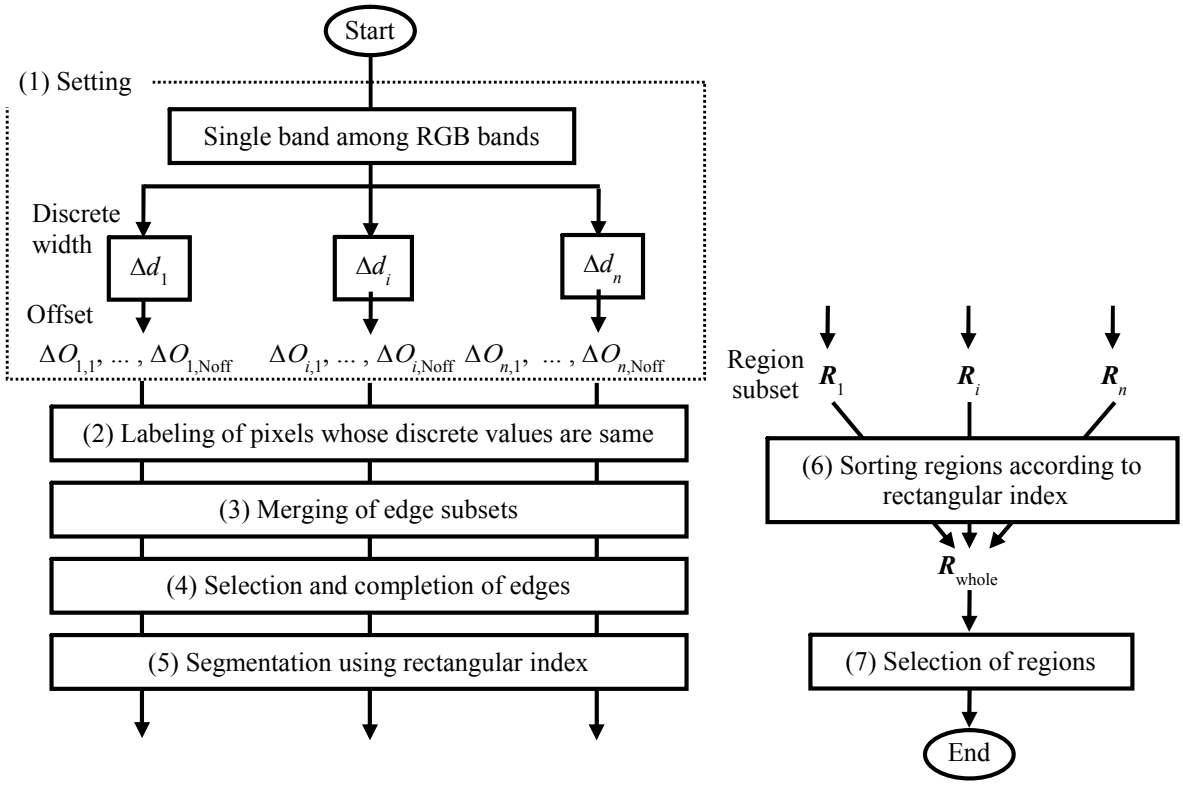


Figure 1: Segmentation flowchart of rectangular regions using image

length of the smaller region and that of the larger region is calculated, and if it is smaller than a threshold, the pair is rejected. Step (4) assumes that two gable-roof buildings are available in reality, but that the four roof areas of these two building are not properly extracted by segmentation of the aerial image. These regions are denoted "group of 1 to 2 roofs" hereafter in this paper. Two pairs of roofs are generated by dividing Region γ into two regions, Region α and Region β (Figure 4 (right)).

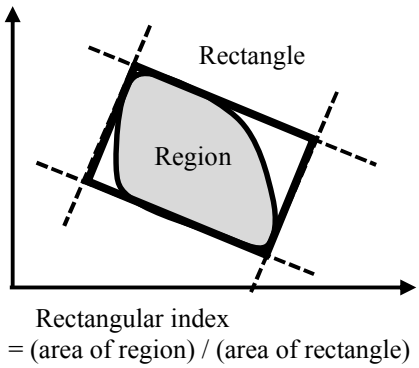


Figure 2: Definition of rectangular index

The proposed algorithm includes three modeling types. Until Step (4), two lists are generated: a list of region pairs, and a list of regions without pairs. Modeling (1) uses the first list. It generates models for those regions whose normals are not vertical and whose partners are available, in descending order of each region's area. A point shared by two regions is determined as a point on a ridgeline. If no point is shared by two regions, a centroid of edges of the two regions is used as an alternative. The ridge direction is established such that it is orthogonal to the azimuthal direction of the two regions normals. A search is conducted, on both sides along the ridge direction, for regions whose normal azimuthal direction is parallel to the ridge's direction. If two regions are detected, a two-sided hip-roof building model is generated. If one region is detected, a single-sided hip-roof building model is generated. If no region is detected, a gable-roof building model is generated. In modeling, the minimum length of the rectangle and minimum area are checked to exclude small models.

Using the list of regions without pairs, Modeling (2) generates models for the regions remaining after Modeling (1). If a region has a mixed normal class and its two azimuthal directions are linearly symmetric (e.g., Region γ of Figure 4), a gable-roof building model is generated. If the region has a single and a slanted normal, a slant-roof building model is generated. Otherwise, a flat-roof building model is generated.

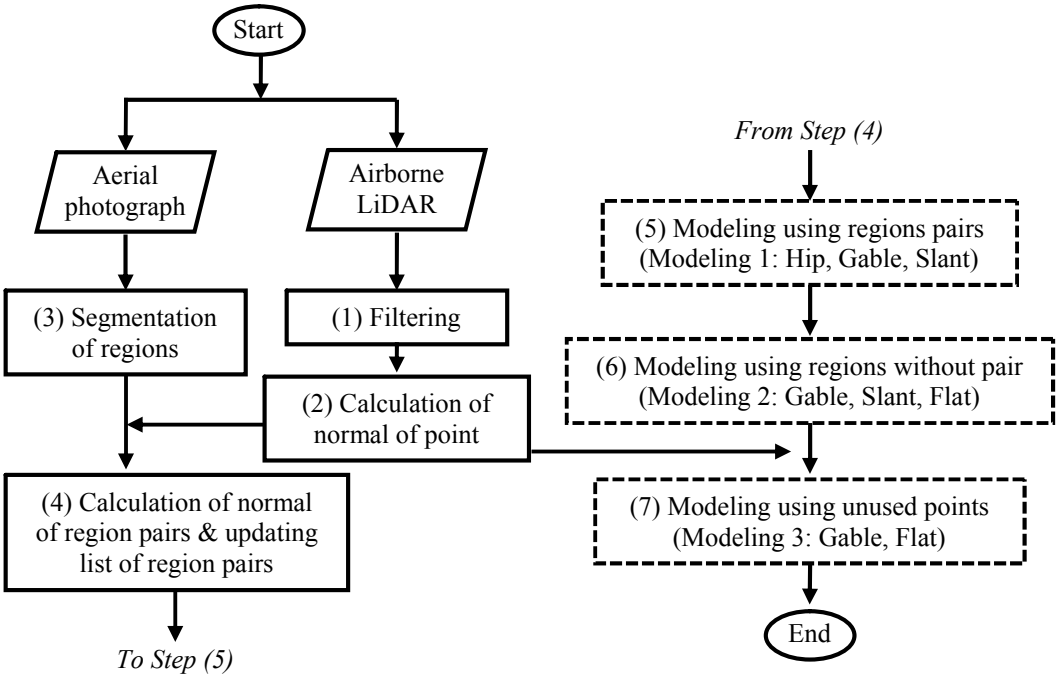


Figure 3: Modeling flowchart using airborne LiDAR and aerial photograph

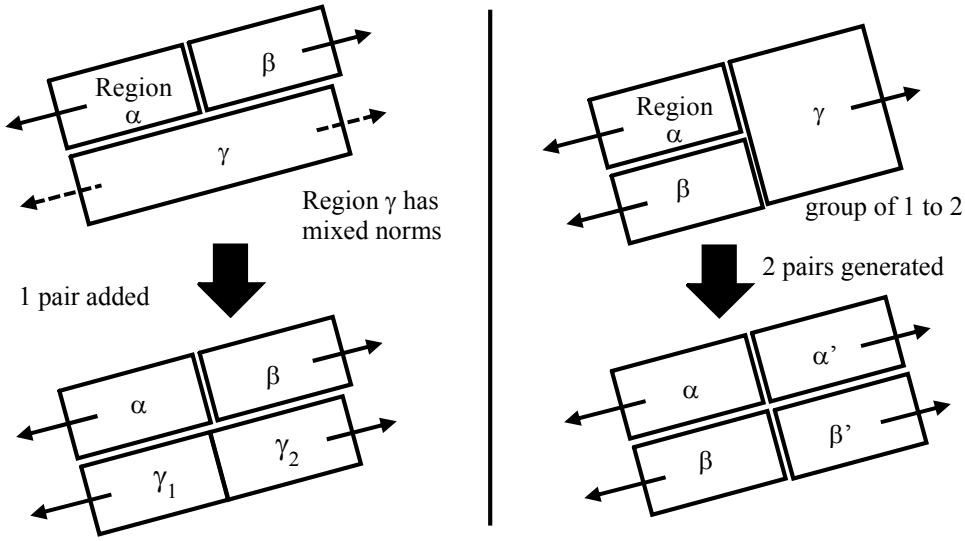


Figure 4: Update of regions. (Left) Region γ is divided into a pair of regions when the distribution of normal indicates two normal that have opposite directions are available. (Right) "Group of 1 to 2" generates two pairs when the normal of Region α and β are opposite to the normal of Region γ.

After Modeling (2), some points may remain unused because of modeling failures. In addition, there may be gaps between the LiDAR data and aerial image observations. When a building did not exist in the aerial image observation but did exist in the LiDAR data observation, the area of the building is not segmented. In this case also, the points of the building area will remain unused. Modeling (3) generates models using the unused points. Similarly to Modeling (2), further gable-roof or flat-roof building models are generated using the azimuthal direction. If some points still remain unused after this process, flat-roof building models are generated without using the azimuthal direction.

In the proposed algorithm, the boundaries of each building are assumed to be orthogonal or parallel to the principal azimuthal direction of its associated group of buildings. The boundaries of a model are delineated so that their

directions are parallel or orthogonal to the principal azimuthal direction of the point clouds in which the model is included. Corrections are performed when the boundaries of models are overlapped.

3. STUDY AREA AND DATA USED

Kyoto is famous for being the old capital city of Japan, and it still has many traditional houses and landscapes. Here, the areas in Higashiyama ward were selected for study. Higashiyama ward is hilly and is well-known for its numerous temples and shrines. Aerial images taken with a 25 cm spatial resolution and orthographically projected were obtained for this research on November 17, 2007. The LiDAR data were obtained in June 2002. The LiDAR density was calculated by finding the ratio between the number of valid pixels and the total pixels for the area. In the case of a 1 m cell size, it was found to be approximately 0.7 points/m^2.

4. RESULTS

First, image segmentation was conducted. Figure 5(a) shows aerial images for a 125 m × 125 m area containing low-rise buildings, segmented by using the algorithm proposed by (Susaki, 2012a). Figure 5(b) to (d) shows segmentation results with Δd_i = 40, 30, and 20. Figure 5(e) shows that the final segmentation results merging the results of Figures 5(b) to (d), and Figure 5(f) shows the segmentation results obtained using ENVI EX 4.8.

Airborne LiDAR data was filtered by using the algorithm proposed by (Susaki, 2012b). Figures 6 and 7 show aerial photographs, segmentation results and modeling results for the two study areas, where the area of Figures 6 (75 m × 75 m area) has low-rise and high-rise buildings and the area of Figure 7 (125 m × 125 m area) has low-rise buildings. The parameters used in the proposed algorithm are shown in the middle column of Table 1.

In this research, generated building models were assessed in terms of types of buildings, the number of buildings, or the ground area of the models. The models generated in the area of Figures 6 and three other areas were validated by comparing the number of models with the actual number of buildings (Table 2). The four areas have 75 m × 75 m area. Then, the models generated in the four areas were validated by comparing the total ground area of the models with that obtained from manual interpretation of the aerial photographs (Table 3). Buildings with actual ground areas of greater than or equal to 30 m^2 were used to perform this validation.

5. DISCUSSION

The proposed algorithm merges regions segmented using the edges detected with different DN interval widths, Δd_i. This quantization can function as a type of spatial filtering, and the algorithm merges the results obtained by smoothing with different spatial scales and. However, the edges are preserved and the scale of spatial filtering is optimized with respect to the size of building roofs in a locality. In the algorithm, regions with a high rectangular index are selected from the regions generated at each quantization. Figure 5 demonstrates that this selection was performed by selecting region candidates among the results obtained using different DN interval widths.

Using the rectangular index, the proposed algorithm extracts regions whose shape is close to being rectangular. However, for roughly textured roofs or in dense urban areas where building boundaries are often unclear, successful detection of complete edges is nontrivial. As shown in Figure 5, the quantization of DNs and the combination of results for several interval widths help to distinguish these roughly-textured roofs and unclear boundaries.

In terms of building modeling, Table 2 shows that the producer′s accuracies for gable-roof, flat-roof and slant-roof buildings were all within acceptable limits, whereas the hip-roof producer′s accuracy was not acceptable. Assessment to separate segmentation accuracy from modeling accuracy was also conducted. Modeling of gable-roof buildings was successful even when using insufficient segmentation results. However, modeling of hip-roof buildings was difficult even with completely successful segmentation. In the unsuccessful cases, the side roofs of the buildings had a low number of LiDAR points and were not of a single normal type. As a result, the side roof was not regarded as a part of the hip-roof building. Especially in dense urban areas, small roofs may not contain enough data points to calculate their normal, owing to occlusion. The results also show that smaller hip-roof buildings may not be correctly modeled because of the incident angle of the LiDAR.

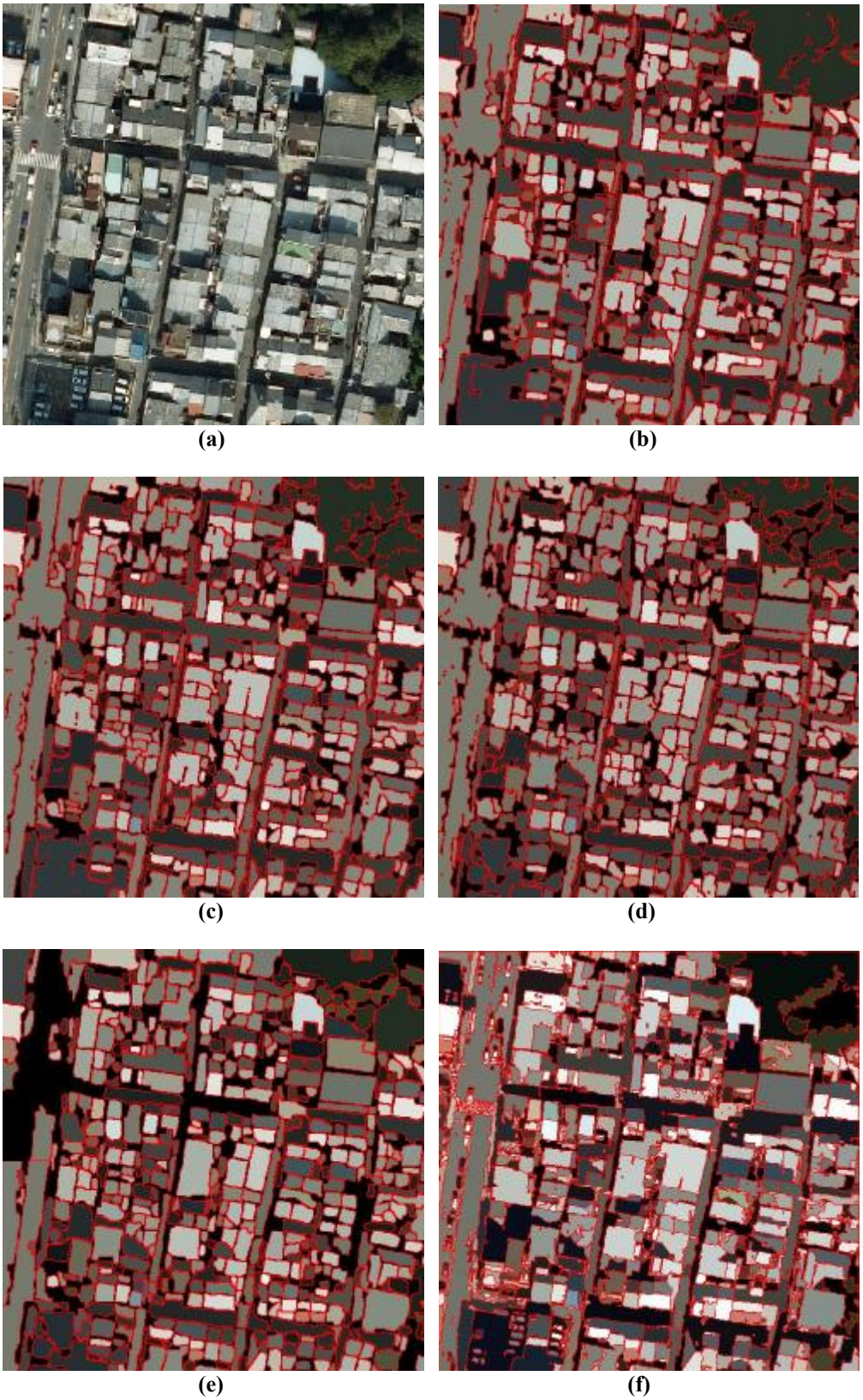


Figure 5: Segmentation of regions: (a) aerial photograph of 125 m × 125 m area, (b) segmentation results with Δd_1 = 40, (c) results with Δd_2 = 30, (d) results with Δd_3 = 20, (e) segmentation results merging (b) to (d), and (f) segmentation results using ENVI EX 4.8. The "feature extraction" function in ENVI EX requires the setting of two parameters, "Scale Level" and "Merge Level", and from an empirical examination, these parameters were set to 50 and 80, respectively.

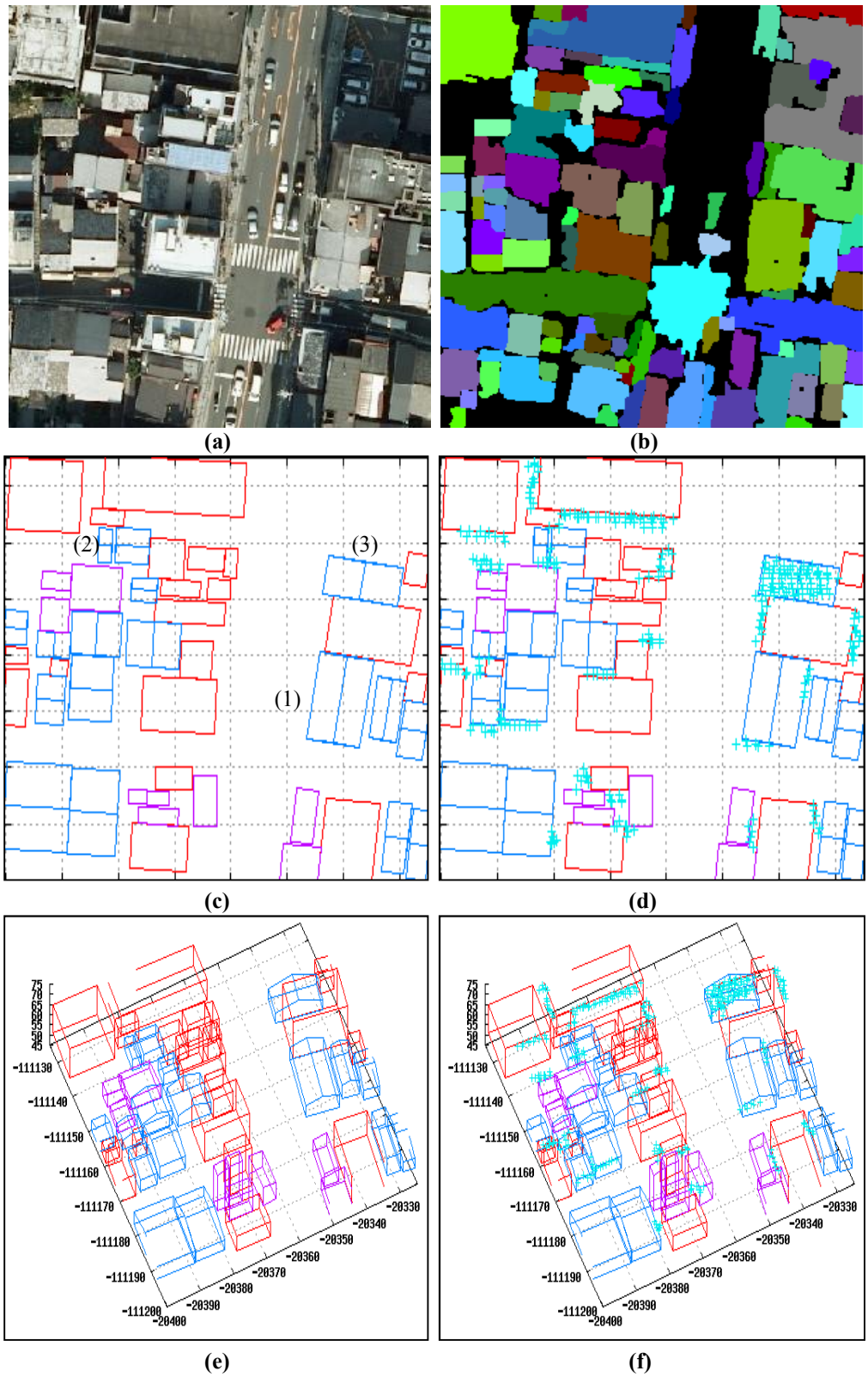
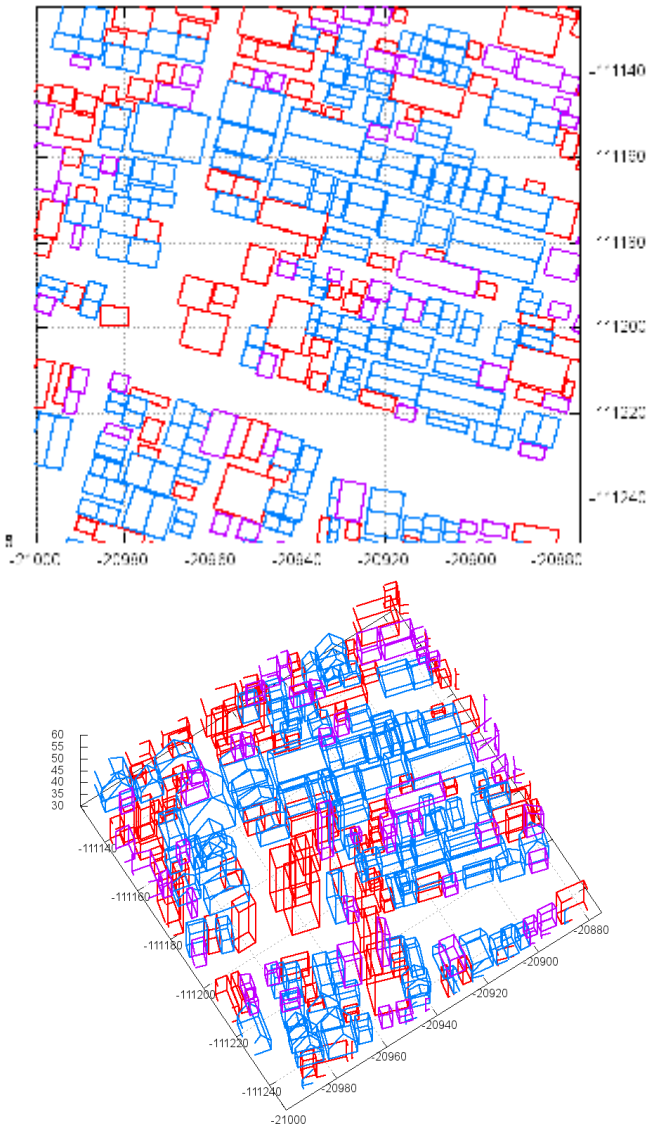


Figure 6: Aerial photograph, segmentation results and modeling results for an area having low-rise and high-rise buildings. (a) aerial photograph has a 75 m × 75 m area, (b) segmented result by using the algorithm proposed by (Susaki, 2012a), (c) wireframe models generated by the proposed algorithm, and (d) the wireframe models incorporating point clouds that remained unused after Step (6). In the models, red denotes a flat-roof building, blue denotes a gable-roof building, pink denotes a slant-roof building and blue crosses denote originally unused points.

(a) (b)



(c) (d)

Figure 7: Aerial photograph, segmentation results and modeling results for an area having low-rise buildings. (a) aerial photograph has a 125 m × 125 m area, (b) segmented result by using the algorithm proposed by (Susaki,

2012a), and (c)(d) wireframe models generated by the proposed algorithm. See Figure 6 for a description of wireframe models.

Table 3 shows that the proposed algorithm underestimated the total ground area of the models by 21%. In dense urban areas, accurate delineation of building boundaries is nontrivial when segmenting regions from aerial photographs. Therefore, boundaries were corrected by checking for overlap between models. As a result, the ground areas of models are expected to be underestimated for dense urban areas.

The proposed algorithm can generate different types of building models. The primary factor in realizing a successful model is best utilization of the region areas provided by the segmentation from aerial photographs. The segmentation algorithm itself has room for improvement. For the case of failing to separate two-roof regions, the proposed algorithm generates a gable-roof building model if the regions possess two normals that are linearly symmetric, and generates two pairs of gable roofs when "group of 1 to 2 roofs" are found in the segmentation results. By employing these approaches, the proposed algorithm utilizes the topology of the buildings for modeling, and effects resulting from segmentation errors are reduced for the proposed algorithm. For example, Area (1) in Figure 6 show generated models when regions possess two normals that are linearly symmetric, and Area (2) in Figure 6 shows generated models when "group of 1 to 2 roofs" are found in the segmentation results.

Table 1: Experimental parameter values

Process	Parameters	Value used
Step (4): Updating list of region pairs	Minimum possibility for single normal class	0.7
	Minimum possibility for two normal classes	0.8
	Minimum valid length	1 m
	Maximum invalid length	1 m
	Minimum ratio between maximum length of the smaller region and that of the larger region	0.25
Modeling (1), (2) and (3)	Minimum length of rectangle	2.0 m
	Minimum area	8.0 m^2

Table 2: Comparison between the buildings contained in the aerial image and those produced by the algorithm. Values in the table denote the number of corresponding buildings.

		Model						
		Gable	Hip	Flat	Slant	Failure	Sum	Producer's accuracy
Actual	Gable	89	0	8	6	0	103	86%
	Hip	2	3	1	0	0	7	50%
	Flat	0	0	25	0	1	26	96%
	Slant	1	0	1	7	0	9	78%
	Sum	92	3	35	13	1	144	
	User's accuracy	97%	100%	71%	54%			86%

Table 3: Absolute and relative error of ground area of building model.

Area (75 m × 75 m area)	Absolute error (m^2)	Relative error (%)
Area 1 (Figure 6)	16.9	20.4
Area 2	25.3	24.7
Area 3	8.0	16.8
Area 4	11.4	20.7
Total	16.2	21.0

Buildings whose regions are not extracted in segmentation can be modeled in the proposed algorithm. Step (7) generates models by using point clouds that are not in currently in use in the model, but that indicate that a building most likely exists. Area (3) in Figure 6 shows that the proposed algorithm succeeds in modeling and compensates for segmentation failure. Although models were not generated in these areas by Step (6), Step (7) did generate correct model types: gable-roof or flat-roof buildings. The processing in Step (7) also handled point clouds reflected from tall trees. An examination of the areas revealed that in estimating the coefficients for the planar equation, a proportion of point clouds reflected from tall trees had small RMSEs. Therefore, the brightness of aerial photographs was utilized to avoid generating false models.

The compensatory factors built into the algorithm are also useful in terms of the difference in time between the acquisition of LiDAR data and aerial photograph observations. Segmentation results are for building data when the aerial photograph was taken. Whether the photograph was taken before or after LiDAR observations, certain buildings that exist in the photograph may not have been captured by LiDAR. Such buildings are removed from the modeling process by counting the number of effective points that have significantly higher elevation than that estimated through filtering. In contrast, if buildings observed by LiDAR are not segmented due to segmentation failure, or are not available on the aerial photograph due to demolition or construction, these buildings can be modeled by performing Step (7). Therefore, effects due to observation delays can be eliminated, and the proposed algorithm can generate models for all buildings available when the LiDAR data are acquired.

6. CONCLUSIONS

In this paper, an algorithm is proposed that automatically generates 3D building models for dense urban areas by using the results of building segmentation from an aerial image. By considering the segmented regions and the roof normals, models are generated of real building types: hip-roof, gable-roof, slant-roof and flat-roof buildings. The proposed algorithm was applied to two study areas of Higashiyama ward, Kyoto, Japan. Owing to the information about the building regions provided by segmentation, the modeling of dense urban areas was successful. Although, on occasion, two roofs were not successfully separated by the segmentation, the proposed algorithm attempts to generate accurate models by referring to the normals in the region. Moreover, the proposed algorithm is able to use point clouds remaining after modeling through segmentation, which indicate that a building feasibly exists. This technique thus compensates for segmentation failure and generates models from LiDAR observations when there is a delay between acquisition of LiDAR data and aerial image observations. The modeling results for the study areas show that it generates relatively simple models when compared with other techniques in the literature, but these models can help users to generate detailed models manually. Performance validation results showed that the proposed algorithm can effectively generate models, especially gable-roof models, even when information from the segmented image was insufficient. High-density LiDAR data can also be used by the proposed algorithm, but the contribution of this extra data to improvement of modeling accuracy and to more detailed modeling might be limited for the case of application to dense urban areas. In the future, roof and side texture should be added to models from aerial images and mobile mapping systems, respectively.

ACKNOWLEDGMENTS

This research was supported by a Grant-in-Aid for Scientific Research (KAKENHI) for Young Scientists (B) (24760412). The author expresses gratitude to Wesco Co. Ltd., Okayama, Japan for providing aerial photographs used for experiments.

7. REFERENCES:

Awrangjeb, M., Ravanbakhsh, M., and Fraser, C. S., 2010. Automatic detection of residential buildings using LIDAR data and multispectral imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 65 (5), pp. 457-467.

Advanced Image Analysis for GIS from ITT. ENVI Feature Extraction Module (ENVI FX), ENVI Feature Extraction Module (ENVI FX); Available online: http://www.ittvis.com/language/en-us/productsservices/envi/enviex.aspx

Kim, K. and Shan, J., 2011. Building roof modeling from airborne laser scanning data based on level set approach. ISPRS Journal of Photogrammetry and Remote Sensing, 66 (4), pp. 484-497.

Lecumberry, F., Pardo, Á., and Sapiro, G., 2010. Simultaneous object classification and segmentation with high-order multiple shape models. IEEE Transactions on Image Processing, 19 (3), pp. 625 - 635.

Rottensteiner, F., Trinder, J., Clode, S., and Kubik, K., 2007. Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis. ISPRS Journal of Photogrammetry and Remote Sensing, 62 (2), pp. 135-149.

Sampath, A. and Shan, J., 2010. Segmentation and reconstruction of polyhedral building roofs from aerial Lidar point clouds. IEEE Transactions on Geoscience and Remote Sensing, 48 (3), Part: 2, pp. 1554 - 1567.

Sohn, G., and Dowman, I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing, 62, pp. 43-63.

Susaki, J., 2012a. Segmentation of shadowed buildings in dense urban areas from aerial images", Remote Sensing. 4, pp. 911-933 doi:10.3390/rs4040911.

Susaki, J., 2012b. Adaptive slope filtering of airborne LiDAR data in urban areas for digital terrain model (DTM) Generation. Remote Sensing, 4, pp. 1804-1819, doi:10.3390/rs4061804.

Tzotsos, A., Karantzalos, K., and Argialas, D., 2011. Object-based image analysis through nonlinear scale-space filtering. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 66 (1), pp. 2-11.