Fusion of LiDAR and Photogrammetric Imagery for Precise and Automatic Building Extraction

Lee DongHyuk † , Lee KyoungMu † and Lee SangUk †

† : School of Electrical Engineering, ASRI, Seoul National University and Image Information Research Center, Korea, odyssey189@diehard.snu.ac.kr, kyoungmu@snu.ac.kr, sanguk@ipl.snu.ac.kr

Abstract We propose a new building detection and description algorithm for LiDAR data and photogrammetric imagery using splitting and merging segments, and line segments matching. Our algorithm consists of three steps. In the first step, we extract initial building regions from LiDAR data. In the second step, given the color segmentation results from the photogrammetric imagery, we extract coarse building boundaries based on the LiDAR results with region segmentation and merging from aerial imagery. In the third step, we extract precise building boundaries based on the coarse building boundaries using line segments matching and the characteristic of man-made object. Experimental results on multisensor data demonstrate that the proposed algorithm produces accurate and reliable results.

1 Introduction

A lot of photogrammetry research has been focused on the development of techniques to reconstruct the boundary representation of building objects in high-density urban areas. The automation of this building extraction technique is essential in dealing with the increased demand for 3D city modelling as an efficient method for compilation of building layers, or as a base layer for 3D reconstruction of building models. In earlier days, most building extraction techniques have relied on 2D feature analysis, in which straight lines or corners extracted from aerial photographs are perceptually grouped, but the building model targeted is rather strongly constrained to decrease the uncertainty of building hypothesis generation and its verification. For achieving more unconstrained building representation, this technique has been extended to 3D feature analysis using multiple imageries, in which higher density of features belonging to building structure, are generated and grouped. To this end, 3D corners, 3D lines, and 3D planar polygons are used. Since the aforementioned approaches rely on the reliability and density of features extracted from aerial photographs, these techniques suffer difficulties when extracted features are highly fragmented or missed due to low contrast, occlusion and shadow effects.

Thus other data sources have been exploited for compensating for the disadvantages of photogrammetric imagery. In this area, LiDAR data, which is acquired by airborne laser scanners, has been used as an attractive alternative to photogrammetric imagery due to high vertical accuracy and high point density. As a single source, LiDAR data has been used to reconstruct various types of building shape - parametric model, prismatic model with flat roof and polyhedral building with the restriction on building orientation. Although LiDAR data have several advantages of building localization and planar patch extraction compared to aerial photographs, there is also drawback to delineate building boundaries with break-lines when LiDAR data is solely used, even with extremely high density of 7 points per square meter. Therefore, an information fusion technique is recommended to combine the complementary nature of the two different data sources for building extraction.

2 System Overview

LiDAR and photogrammetric imagery each have particular advantages and disadvantages in horizontal and vertical position accuracy. Compared with photogrammetric imagery, LiDAR provides accurate height information but inaccurate boundary lines. Unfortunately, some regions in LiDAR data have null values due to self-occlusion of a building. Photogrammetric imagery provides extensive 2D information such as high-resolution texture and color information. Although 3D height information can be estimated from one or several images by several methods such as stereo, shape from shading compared with

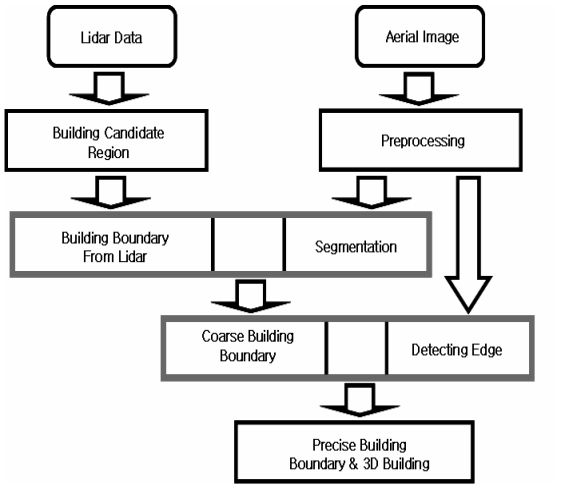


Fig. 1. Block diagram of the system.

LiDAR, this information is relatively inaccurate.

To extract accurate building regions, we combine the information from LiDAR and photogrammetric imagery by first assuming the building regions extracted from LiDAR as the initial regions, and then improving the inaccurate boundaries by utilizing the photogrammetric imagery to obtain the precise building boundaries. More specifically, we first extract the initial building boundaries from the height information provided by LiDAR data. Next, we enhance the initial boundaries using color information provided by the photogrammetric imagery. Here, we apply color segmentation based on the assumption that building roofs are planar and therefore the height of the building and its color depicted in the photogrammetric imagery is homogeneous. Finally, we apply edge matching and closed loop construction as a post process, since most man-made objects are best described by edges. Figure 1 shows the block diagram of the system.

3 Extracting Coarse Building Boundaries

Any automatic or manual point-matching method outputs some error in the processing of registering of data captured by photogrammetric and LiDAR systems to a common reference frame. Moreover, hills and trees around buildings or height variations in some buildings can disturb accurate extraction of building boundaries from LiDAR data. Therefore, before extracting precise building region, coarse building regions are extracted from the photogrammetric imagery via a over-segmentation and segment merging technique together with the building extraction results obtained based on the LiDAR data from the previous process.

A. Registration

In this paper, we use high resolution aerial images as photogrammetric imagery and assume that there are no internal/external orientation parameters available for the aerial image and LiDAR data. Therefore, we estimate homography for the registration of LiDAR and aerial image [4].

B. Color segmentation

Most conventional color segmentation algorithms are based on the information in the color space. However, clustering in color space only usually does not provide satisfactory performance, since it lacks information about the spatial configuration. To resolve this limitation, in addition to color, the spatial coordinates of pixels are often incorporated into their feature space representation. Here, we apply the mean-shift color segmentation method [5] to obtain over-split segmentation results. As a result of this segmentation, one building region is divided into one or more segments. Figure 2(a) shows an example of the over-segmentation results.

C. Extracting a coarse building region

A merging algorithm to merge separate segments of a single building is required to extract building regions based on the over-segmented results of the aerial images. Our merging method is based on the preliminary building regions and their heights obtained from the LiDAR data in previous stages. We define the ***support ratio*** as the ratio between the area of a particular segment that is supported by the LiDAR data to its total area as follows:

$$ratio = \frac{L_a \cap S_a}{S_a}$$

Where L_a is the area of the building region extracted from LiDAR data, S_a is the area of a segment in aerial image. The *support ratio* naturally incorporates information extracted from the LiDAR data to represent the probability that a particular segment is a building region. If the ratio of a segment is more than a predefined threshold, then the segment becomes a building region. Finally, segments that are classified as building regions by common building regions in LiDAR data are merged into a group. In addition to the split and merge routine, color data from aerial images are used to remove tree segments directly connected to building regions, since they distract accurate building boundary extraction. More specifically, relatively small segments (in the image) in a building region (which is determined in the LiDAR data) that have high green value are removed. As shown in Figure 2, 4 the partitioned building regions in the middle of Figure 2(a) are connected in Figure 4(a) by this split and merge method. Moreover, the majority of the tortuous building boundaries in Figure 2(a) are trimmed in

(a) (b)

Fig. 2. Extracting coarse building region. (a) Segmentation result by mean-shift method, (b) Detected building region from LiDAR data

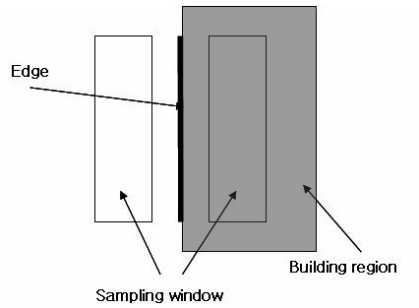


Fig. 3. Building boundary candidate verification with height information

Figure 4(a) and tree segments directed connected to building regions are removed.

4 Extracting Precise Building Boundaries

Although it is possible to correct building regions from LiDAR data by the coarse building boundary process, the precise building boundaries cannot be obtained, since just trimming the tortuous coarse building boundaries may not guarantee the accuracy of extracting building boundaries process. The precise building boundaries are obtained in a closed-loop form by extracting edges from an aerial image, selecting the edge corresponding to the coarse building boundary, and conducting post-processing by perceptual grouping such as edge linking and closed loop constructing.

A. Building boundary trimming

When applying edge information from the aerial image, it is difficult to use the coarse building boundary as the initial input. In this paper, to find the corner points of a building boundary, the Douglas-Peuker method is applied. This method is a recursive split method to find the trimmed coarse building boundaries.

B. Extracting building boundary candidates

As a feature in aerial image, the edge that is the best descriptor of the building is used. To extract the edge, the Nevatia-Babu edge detector is used. From detected edges, not all the edges are used.

Only edges near a building boundary are necessary for the precise building boundary. Therefore, we can reduce the number of edges (building boundary candidates) by using information from the coarse building boundary and heights. Restricting the number of building boundary candidates reduces the computing cost and allows us to extract reliable building boundary candidates.

B.1 Searching for edges around a building

An easy and simple method of restricting the number of boundary candidates is to find the edges near the building boundary. We accept edges within 20 pixels of a boundary as boundary candidates, considering registration error.

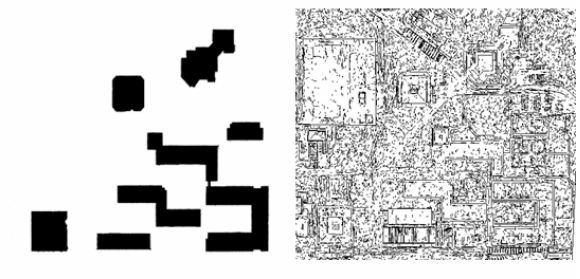
B.2 Height compatibility

Height data from LiDAR measurements are used to eliminate some of the non-boundary edges around a building. If the registration of LiDAR data and photogrammetric imagery is correct, as shown in Figure 3, an edge in photogrammetric imagery is verified as a boundary candidate when its height is compatible with the average heights of the sampling window placed at either side of that edge. Note that the building inside region of a flat or sloping roof is higher than the building outside region of the roof boundary. Therefore, if an edge is part of a building boundary, then there is greater height difference between average heights of the two side sampling windows than the difference from bare-earth edge. In this paper, after extracting building candidate edges, binary values 0 and 255 have been assigned to building regions and bare-earth regions, respectively. If an edge is building boundary, then the difference of average value between left and right sampling window should be about 255. In other words, as shown in Figure 3, if an edge is building boundary edge and if the average value of the left sampling window is 255 (bare earth region), then the average value of right sampling window should be 0 (building region), and vice versa. Using this attribute, we can select building boundary candidates from edges around a building.

C. Constructing precise building boundaries

C.1 Line segments matching

Using the extracted building boundary candidates described above, a boundary in the coarse building boundary result can be paired to several edges in the building boundary candidates. Several attributes exist for resolving matching ambiguities. The idea is to restrict the candidate matches to match one-to-one by some constraints, such as length ratio, angle or distance between coarse building boundary and building boundary candidates. Because if we allow a 1:n relationship,



(a) (b)

Fig. 4. Extracting precise building region. (a) Coarse building region extracted from Fig. 2, (b) Detected edge by Nevatia-Babu edge detector

shadow or street edges can disturb the performance of matching. For a given pair of line segments to be considered as a candidate match, constraints on the three attributes are applied to select the most consistent matches. First, the length ratio of a match can eliminate very short line segments that are false matches. Second, the distance measure can eliminate far-away edges. In this paper, we define the distance of two edges as the average distance between a point on one edge and a point on the other edge line. The distance between l_1 and l_2 can have two cases. Considering this, the distance is defined as follows:

$$E_d = ((\sum_i^n d_{1i})/n + (\sum_j^m d_{2j})/m)/2$$

where d_{1i} is the distance between l_1 edge and a point on l_2 edge, while d_{2j} is the distance between a point in l_2 edge and l_1. Due to distance error in the registration process mainly caused by point matching error, we allowed a maximum of 20 pixels distance between the coarse building boundary edge and building boundary candidate edges. Third, the similarity angle of two line segments, which is the most important attribute for defining similarity in relative position, is measured and pairs with angles over a threshold are discarded. We allowed a maximum of 15 degrees between two line segments. This process is applied to each building group. The combination of these constraints can be represented as follows:

$$Score = \begin{cases} w_1 L_R + w_2 \dfrac{15-\theta}{15} + w_3 \dfrac{20-E_D}{20}, & \text{if } \theta < 15 \text{ and } E_D < 20 \\ 0 & \text{otherwise} \end{cases}$$

where θ denotes the angle between the two line segments, E_D denotes the distance between the two line segments, and w_1, w_2, w_3 denote weights of each attribute. L_R is the ratio of the two lengths given by:

$$L_R = \begin{cases} \dfrac{C_L}{E_L} & \text{if } E_L > C_L \\ \dfrac{E_L}{C_L} & \text{if } C_L > E_L \end{cases}$$

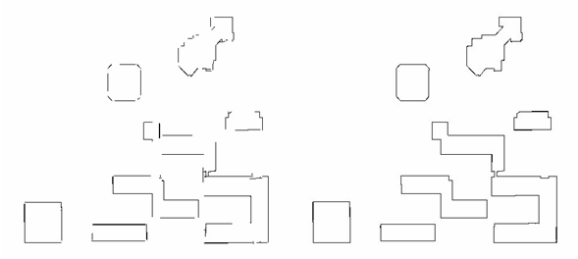
where C_L is the length of the coarse building boundary and E_L is the length of the building boundary candidate. In our implementation, the weights $w_1 = 1$, $w_2 = 2$, $w_3 = 3$ and a threshold score of 3.0 is used.

It should be noted that no corresponding match exists when the score value is lower than the threshold. There are two strategies for this case. The first is to substitute an unmatched building boundary candidate with the corresponding coarse building boundary. This has the disadvantage of using inaccurate line segments but guarantees minimum accuracy. The second is to skip the unmatched building boundary. In other words, in this strategy, only matched edges are used to construct building boundaries. The advantage of this is that an edge in the building boundary candidates can compensate for a collapsed or false initial building boundary, but the accuracy of this strategy can be poor if all building boundary edges are not extracted from the aerial image. If the above constraint is fulfilled, we say that the pair of line segments considered is self-consistent and can form a precise building boundary. For each line segment in the coarse building boundary, we thus have a line segment match from building boundary candidates. Figure 5(a) shows the results of line segments matching.

C.2 Construction of a closed loop

The result shown in Figure 5(a) is not the complete building boundary. Since a building is a polyhedral structure of closed-loop form, the building boundaries in this figure should be closed polygons. To satisfy this condition, we must consider the missing features in the building sides and corners. Grouping only neighboring line segments is sufficient to compensate for the missing features. The reason is that the coarse building boundaries have the closed-loop form, and each matched line segment in the building boundary candidates has its one-to-one corresponding coarse building boundary.

Synthetically, we suggest two simple forms of closures. One is applied when the junction angle is more than 45 degrees, and the other is applied when the angle is less than 45 degrees. When the angle is more than 45 degrees, we can consider the intersection point of the extended lines of two edges as the junction point and the lines are extended to that point, respectively. When the angle is less than 45 degrees, the intersection point can be placed in wrong locations.



(a) (b)

Fig. 5. The result of precise building boundaries. (a) Extracted corresponding edges, (b) Closed loop form of (a)

Therefore, in this case, we just link the two edges. Figure 5(b) shows closed loop form of Figure 5(a) by our method.

5 Experimental Results

Experiments designed to demonstrate the performance of our algorithm were carried out on two data sets. We show results for several examples in this section.

A. The format of the input data

The LiDAR data and aerial images used in this paper, obtained airborne, represent a large $2km \times 2km$ area of Daejeon in the Republic of Korea. The laser point density is about $3.2\ pt/m^2$. The resolution of the aerial images is about $5.2\ pixel/m$. Here, in the registration process, we scaled the LiDAR data result up to fit the high-resolution aerial images. This is to utilize all available image data resources.

B. Analysis of extracted building boundary

Unfortunately, it is difficult to acquire large data sets for a valid statistical evaluation. In addition, most of the building detection and description systems have different representational powers, and statistical evaluation on a small number of examples is less meaningful when the results depend on the choice of test data set. In this paper, we use two data sets to analyze the extracted building boundary. Set 1 shows a wide and complex building region and set 2 shows a region where buildings are described in detail and the building wall edges may interfere the line segment matching. Moreover, we compare the accuracy of building boundary extraction results obtained from only LiDAR data without any refinement, and LiDAR data with refinement, and from combining LiDAR data and photogrammetric imagery with intermediate coarse boundary extraction and precise refinement. The refinement process for the results extracted from the LiDAR data case is identical to the precise refinement of the precise refinement of the combined case. To evaluate the results of our proposed method, ground truth of building boundaries are extracted manually from warped aerial images. For quantitative analysis of our result, we selected the data (aerial images) with higher accuracy to use as the ground truth. It is then possible to compare building extraction results with the ground truth. For performance evaluation, Chamfer distance is applied to evaluate each method. To estimate Chamfer distance, the extracted building boundaries are converted from simple edges to a distance image. The result of each method is superimposed on the distance image of the ground truth, and the average of the distance values where the edges of the results intersect with the distance image represents the Chamfer distance. If the edges of a building boundary extraction results fits the distance image of the ground truth perfectly, the Chamfer distance is 0. In this paper, root mean square Chamfer distance is chosen:

$$Ch_{dis} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \mathrm{v}_\mathrm{i}^2}$$

where Ch_{dis} is Chamfer distance, v_i is distance value and n is the number of points.

Table I shows the accuracy of proposed methods. In this table, because we scaled the LiDAR data result up to fit the high-resolution aerial images, 5.2 pixel distance denotes 1 m. Here, a performance improvement of about 1-4 pixel distance was obtained in each step. However, the result of the coarse and precise building boundary in set 2 did not show any improvement in the performance due to error in the segmentation result. This false segment can occur where the texture of the building roof and walls are similar, and this has an effect on the Chamfer distance considerably. This is a limitation of applying the coarse building boundary extraction method based on color segmentation techniques exclusively, but

Table I. Evaluation by Chamfer distance

Method	Data Set	
	Set 1	Set 2
Only LiDAR (pixel distance (meter))	5.56(1.07m)	7.46(1.44m)
LiDAR + refinement (pixel distance (meter))	4.55(0.88m)	3.24(0.62m)
LiDAR + image + refinement(pixel distance (meter))	2.27(0.44m)	3.54(0.68m)

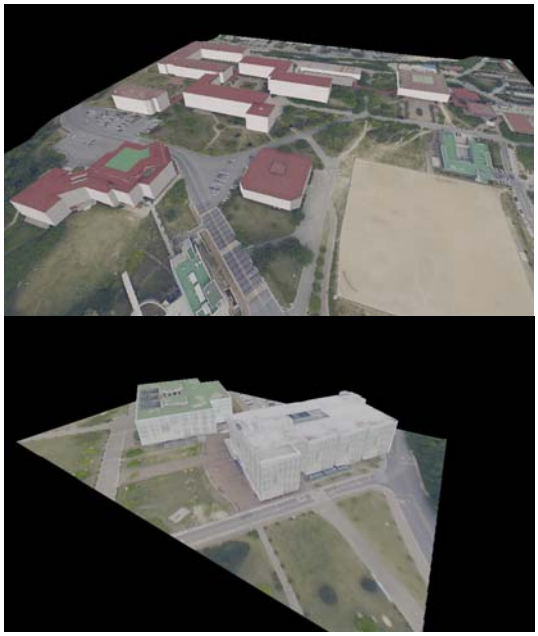


Fig. 6. Reconstructed building results of set 1, 2.

commonly a more detailed description of building boundaries comparing to LiDAR results is obtained by coarse boundary processes.

C. Reconstruction of 3D building

To reconstruct a 3D building model, several techniques are applied. The ground plane is extracted bya mathematical morphology technique, and texture mapping is applied to the meshed DTM and building regions. 3D building regions are displayed from a perspective view to produce realistic visual effects. Figures 6 shows the 3D reconstructed building regions of set 1 and set 2.

6 Conclusion

A new approach to extract the boundaries of complex buildings from LiDAR and photogrammetric imagery has been developed. The method is based on the application of information fusion to compensate for each sensor's shortcomings. We used the Chamfer distance to evaluate the proposed methods. From the evaluation, we have shown that the proposed multisensor fusion-based building detection method has improved the performance substantially in the accuracy of building boundaries compared to those of methods using LiDAR data only. The building boundary accuracy has improved more than 50%, and we could achieve satisfactory 3D reconstruction result based on extracted building boundary.

In this paper, we assume that any registration was not performed between two sensors (LiDAR and photogrammetric imagery) and there are no internal/external orientation parameters available for the registration. To overcome this, we allowed a maximum of 20 pixels distance between the coarse building boundary edge and building boundary candidate edges. The contribution of this algorithm is that color information in photogrammetric imagery is used to complement collapsed building boundaries obtained by LiDAR. Moreover, edge matching and construction of closed roof is used to reflect the characteristic of man-made object.

In our proposed method, many parts of proposed processes are performed separately, since each process in large area data requires high cost. However, in the very near future, the improvement of computing system will enable fast and incorporated processes within reasonable bounds. Therefore, we expect that the use of such methods will be automatically helpful in many other building detection and reconstruction problem domains as well.

Acknowledgment This research was supported by the Defense Acquisition Program Administration and Agency for Defense Development, Korea, through the Image Information Research Center under the contract UD070007AD.
This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Conter) Support program supervised by the IITA (Institute of Information Technology advancement) (IITA-2008-C1090-0801-0018)

References

[1] Rottensteiner, F., and C. Briese, 2002. A new method for building extraction in urban areas from high-resolution Lidar data, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIV, pp. 295-301.

[2] Kim, Z., and R. Nevatia, Automatic description of complex building from multiple images, Computer Vision and Image Understanding, vol. 96, pp. 60-95, 2004.

[3] Sohn, G., and I. Dowman, Building Extraction using LiDAR DEMs and IKONOS Images, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 3/W13, Dresden, Germany, 2003.

[4] Hartley, R., and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2nd edition, pp. 89-92, 2003

[5] Comaniciu, D., and P. Meer, Mean Shift: A robust approach toward feature space analysis, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 1-18, 2002.