Article

Segmentation of Shadowed Buildings in Dense Urban Areas from Aerial Photographs

Junichi Susaki

Graduate School of Engineering, Kyoto Univeristy, Kyotodaigaku Katsura, Nishikyo-ku, Kyoto 615-8540, Japan; E-Mail: susaki.junichi.3r@kyoto-u.ac.jp; Tel.: +81-75-383-3300; Fax: +81-75-383-3300

Abstract: Segmentation of buildings in urban areas, especially dense urban areas, by using remotely sensed images is highly desirable. However, segmentation results obtained by using existing algorithms are unsatisfactory because of the unclear boundaries between buildings and the shadows cast by neighboring buildings. In this paper, an algorithm is proposed that successfully segments buildings from aerial photographs, including shadowed buildings in dense urban areas. To handle roofs having rough textures, digital numbers (DNs) are quantized into several quantum values. Quantization using several interval widths is applied during segmentation, and for each quantization, areas with homogeneous values are labeled in an image. Edges determined from the homogeneous areas obtained at each quantization are then merged, and frequently observed edges are extracted. By using a "rectangular index", regions whose shapes are close to being rectangular are thus selected as buildings. Experimental results show that the proposed algorithm generates more practical segmentation results than an existing algorithm does. Therefore, 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. By utilizing these factors, the proposed algorithm optimizes the spatial filtering scale with respect to the size of building roofs in a locality. The proposed algorithm is considered to be useful for conducting building segmentation for various purposes.

Keywords: segmentation; urban; shadowed buildings

1. Introduction

Three-dimensional (3D) modeling of buildings in urban areas has recently gained widespread popularity and has been studied by many researchers. Airborne light detection and ranging (LiDAR) is considered useful to provide cloud points having 3D coordinates and to help in delineating building boundaries. However, as a result of more than a decade of research, for 3D modeling, it has been found to be highly effective to fuse airborne LiDAR data with data from other sources, for example, digital maps [1,2] or remotely sensed images [3–6]. Whereas a digital map is costly, aerial photographs and satellite images are relatively cheap and widely applicable to many areas. Two-dimensional (2D) boundaries of buildings obtained through image classification would aid in creating accurate and effective 3D models.

Classification of remotely sensed images is roughly divided into pixel- and object-based approaches. Pixel-based approaches, for example, clustering, the maximum likelihood method, and Support Vector Machines (SVM), assign class labels to pixels by calculating the probability that a pixel belongs to each class [7]. In contrast, model- or object-based approaches utilize context information from neighboring pixels. One of the most well-known, object-based approaches is to use mathematical morphological classifiers [8–12], and object-based approaches have been applied to segment urban landscapes [13–15]. In general, object-based approaches generate classification results with high accuracy, whereas pixel-based approaches often have 'salt-and-pepper' noise because they assume that the data of each pixel are independent.

The author's particular interest is in dense urban areas, in which houses and buildings are located close to one another and narrow streets are found. The proximity of the buildings causes two problems. First, the boundaries between the buildings are unclear, and second, many shadows are cast by other buildings in comparison with typical urban areas. In addition, as shown in Figure 1, traditional Japanese houses often have undulating slate roofs with a rough texture, and thus the standard deviation of their digital number (DN) is large. This rough texture also causes a third problem, which is that many erroneous edges are detected during segmentation preprocessing. Owing to these features, segmentation results were poor for the area in Figure 1 using an existing algorithm.

The first and third problems can be regarded as being equivalent: the problem is solved by the provision of appropriate edge detection. Canny [16] proposed an edge detection operator that is robust to noise, and this operator is widely utilized. Other edge detection operators, based on wavelet [17,18], multiscale [19–22], and multiscale with Markov random field [23–26] approaches, have also been proposed. Furthermore, algorithms to compensate for the lack of brightness in shadow regions have been presented [27–29]. However, the compensated results reported in [30] showed that the boundaries between originally shadowed and unshadowed regions remained clear, and that over-compensation is an issue yet to be solved.

In this paper, an algorithm is proposed that segments buildings, including shadowed buildings, in dense urban areas from aerial photographs. The data used in this research and the areas under study are described in Section 2. The proposed segmentation algorithm is outlined in Section 3, and experimental results are reported at the end of this section. The algorithm and the experimental results are then discussed in detail in Section 4, and Section 5 concludes the paper.

2. Study Area

Kyoto is the historic capital city of Japan, and still maintains many traditional houses. Areas in Kyoto's hilly Higashiyama ward, which is famous for its numerous old temples and shrines, were selected for this study because they are good examples of dense urban areas in Japan to examine the performance of the segmentation algorithm. The targeted areas have narrow streets, approximately 5–6 m in width. Figure 1 shows an example of the buildings in Higashiyama ward. Orthographically projected, RGB bands of aerial photographs of these areas with a 25-cm spatial resolution were available for this research. The photographs were taken using Ultra CamX (UCX), Vexel.

Figure 1. Houses in the study area (Higashiyama ward, Kyoto).

3. Segmentation Algorithm

This paper focuses on an algorithm to segment buildings from aerial photographs of dense urban areas. As mentioned above, the segmentation of buildings in dense urban areas has a number of difficulties. Here, to distinguish roofs having rough textures, DN intervals are quantized into a number of quantum values, following a similar approach to Deng and Manjunath [31]. Quantization using several DN interval widths is applied during the segmentation algorithm, and for each quantization, areas with homogeneous quantum values are labeled in an image. Edges determined from the homogeneous areas obtained at each quantization are subsequently merged, and frequently observed edges are extracted. Roofs and buildings are then segmented using these extracted edges.

The proposed segmentation algorithm consists of the following steps (see Figure 2). The algorithm assumes that images consist of 1-byte pixels in each of the three color bands (RGB).

1 Set the number of DN intervals for quantization N_{disc}, the associated interval widths Δd_i ($i = 1, \cdots, N_{\text{disc}}$), and the number of offsets N_{off}. The offset width Δoff_i is defined as $\Delta\text{off}_i = \Delta d_i/N_{\text{off}}$, and N_{off} quantized images are generated at a given value of Δd_i by applying the different offsets. For example, with $\Delta d_i = 40$ and $N_{\text{off}} = 5$, Δoff_i is $40/5 = 8$, and the offsets are {0, 8, 16, 24, 32}. With offset = 0, DNs are quantized into the intervals [0, 39], [40, 79], [80, 119], [120, 159], [160, 199], [200, 239], and [240, 255], and all pixels having a DN within an interval are assigned the same

quantum value.

2 Taking each quantized image in turn, regions are extracted and labeled by examining both the four neighboring pixels surrounding a given pixel and all other connecting pixels having the same quantum value. Large regions are removed, and then small regions are merged with neighboring larger regions, if such larger regions exist; otherwise, the small regions are removed. Finally, the edges of any remaining regions are extracted.

3 All edges of the N_{off} quantized images at a given value of Δd_i are merged, and the number of edge detections within each pixel is counted.

4 A pixel whose edge count is greater than or equal to a threshold T_{count1} is preserved as an edge. Moreover, a pixel whose edge count is smaller than T_{count1}, but greater than or equal to T_{count2}, is added to an edge group if the pixel is connected to preserved edge pixels. Finally, a non-edge pixel is changed into an edge if linear alignments of edges pixels are found either side of it.

5 Segmented regions are generated using the edges found in each quantization. To perform segmentation, a "rectangular index" is calculated as follows (see Figure 3).

(1) By using the 2D coordinates of the edges in a region, a main axis and sub-axis are determined, where the sub-axis is orthogonal to the main axis.

(2) The region is then projected onto the main axis, and the maximum, $V_{1,max}$, and minimum, $V_{1,min}$, coordinate values along the main axis are obtained. In the same manner, the maximum, $V_{2,max}$, and minimum, $V_{2,min}$, coordinate values along the sub-axis are obtained. A rectangular area is calculated by using the formula $S_{rect} = (V_{1,\,max} - V_{1,\,min} + 1) * (V_{2,\,max} - V_{2,\,min} + 1)$.

(3) The rectangular index *idx* is defined as the ratio between the actual area of the region S_{actual} and S_{rect},

$$idx = S_{actual} / S_{rect}. \tag{1}$$

Therefore, *idx* ranges from 0 to 1, and a region whose rectangular index is close to 1 has a shape similar to a rectangle.

(4) If *idx* is lower than a given threshold, the region is removed because a strong likelihood exists that the region does not correspond to a building.

6 Regions obtained in the N_{disc} images are sorted according to their rectangular index.

7 Regions with high rectangular index are selected as buildings, as long as no part of the region overlaps with regions already selected. The unselected regions are next considered, and a region is examined if both its overlap area with previously selected regions and the ratio between this area and the region's total area are less than or equal to given thresholds. If *idx* for the portion of the region without overlap is greater than or equal to a further threshold, that portion is added to the group of regions nominated as buildings. Finally, any holes in the buildings are filled.

Some of the steps require more detail. In Step 4, if a target pixel is not an edge pixel, then the numbers of edge pixels in neighborhoods around the target pixel are counted. Figure 4 illustrates the filters used for finding edge pixels in the top-to-bottom, left-to-right, upper-left-to-lower-right, and lower-left-to-upper-right directions. By designating edge pixels as having a value of 1 and non-edge pixels as 0, a score is calculated by multiplying the filter components by each pixel's value. The target pixel is labeled as an edge pixel, if the following conditions are satisfied when applying any filter:

(1) The local scores in Figure 4(a,b) are greater than or equal to T_{count3}.

(2) The total score of all (7 × 7 pixels) components is greater than or equal to T_{count4}.

The second condition prevents mislabeling of non-edge pixels near the corners of the rectangles. The above search is repeated a maximum four times using four different filters.

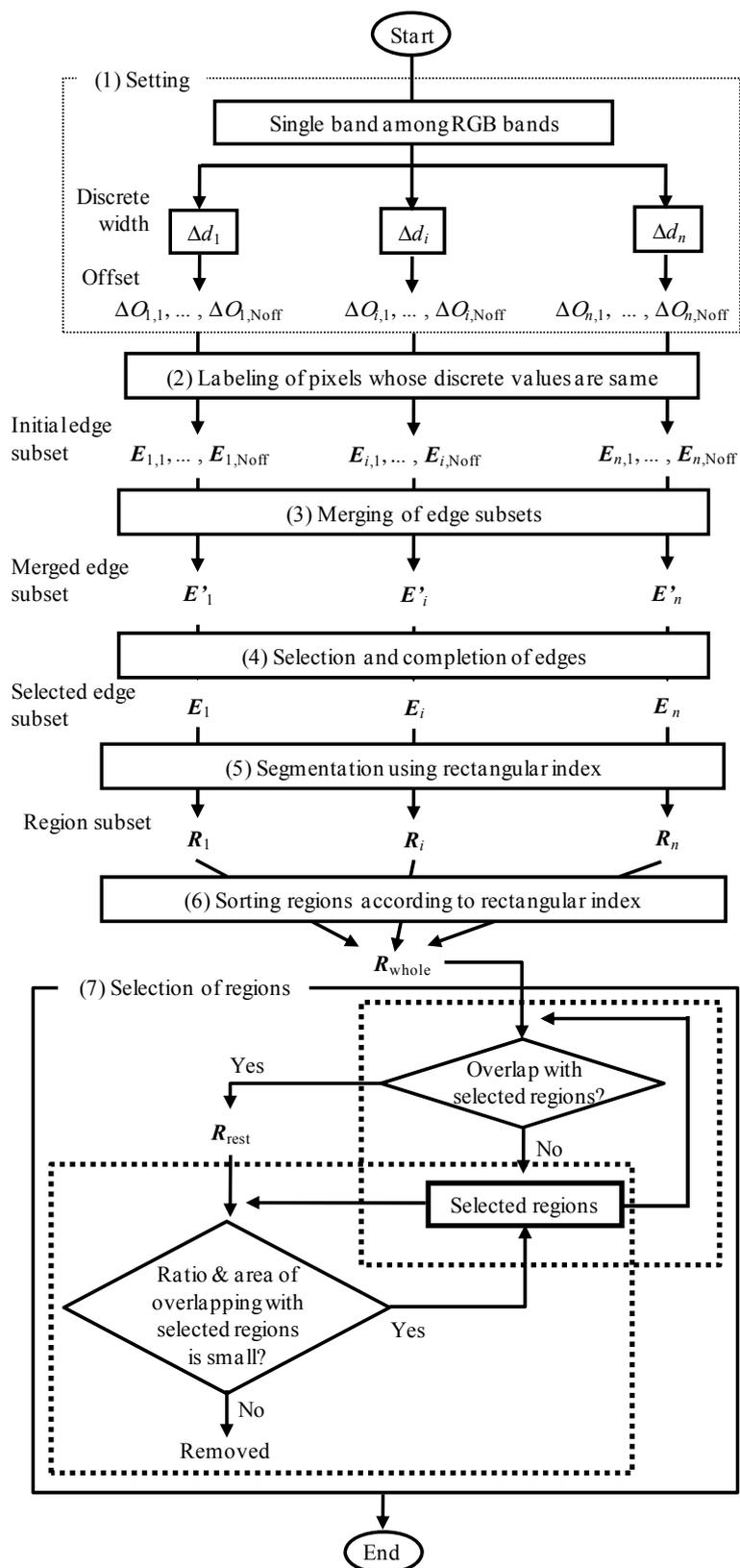**Figure 2.** Flowchart of the proposed segmentation algorithm.
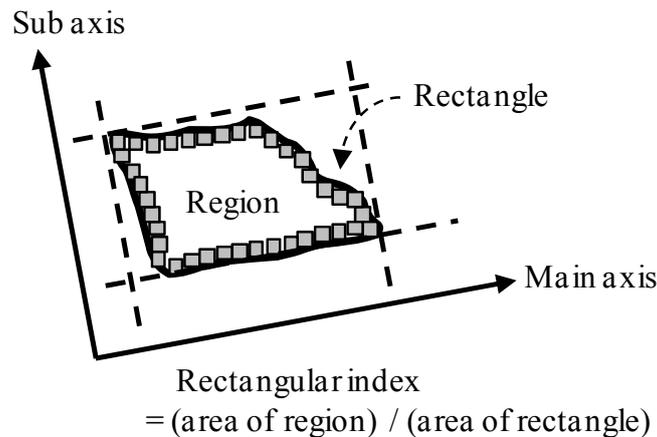
Figure 3. Calculation of the rectangular index.

Rectangular index
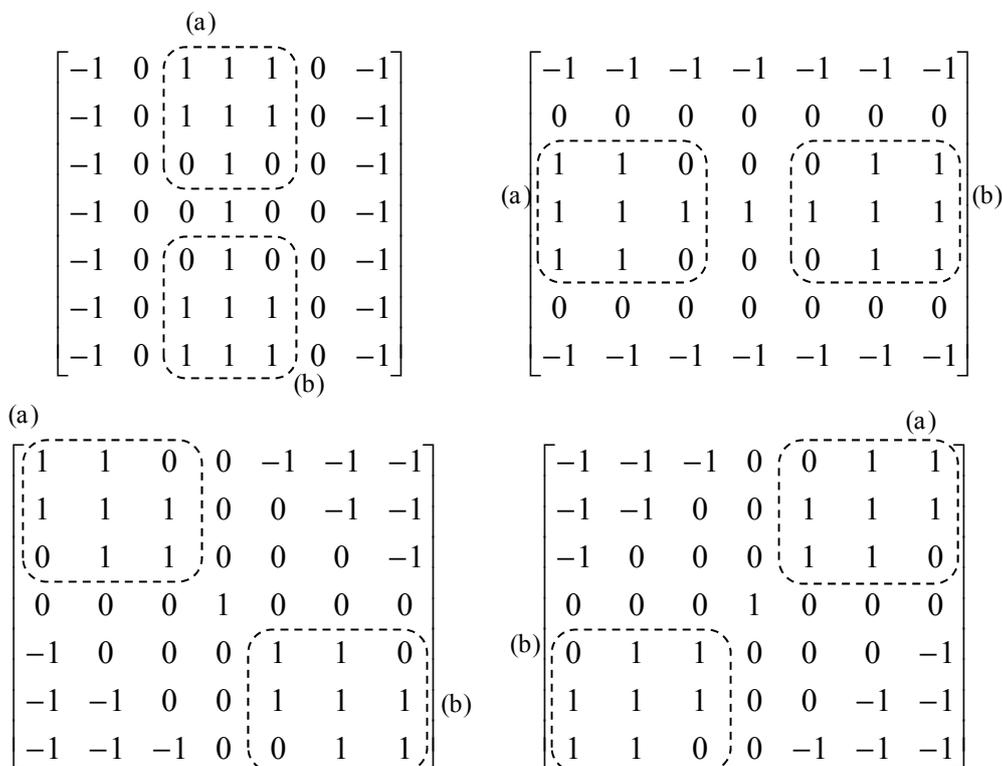= (area of region) / (area of rectangle)

Figure 4. Filters for edge completion.

Finally, calculation of the rectangular index in Step 5 should be clarified. In the algorithm, the main and sub axes are not determined by principal component analysis (PCA). (The reason for not using PCA is discussed in Section 4). Instead, a pair of edges whose distance is within a certain range (d_{edge_min}, d_{edge_max}) is selected, and the angles of the lines connecting the edges are voted. The angle achieving the maximum voting score is selected as the direction of the main axis. The sub-axis is then determined from the requirement that it must be orthogonal to the main axis.

4. Results

In the experiment, the parameters required by the proposed algorithm were set to the values shown in the right-hand column of Table 1. The optimal values of the parameters may depend on the study

area, and they were set empirically through manual checking of the segmented results. Figure 5 shows the result of each step of the algorithm flowchart in Figure 2. Specifically, the results are shown of labeling using the quantized images, edge detection, segmentation, and selection of regions. Three study areas were selected to examine the performance of the proposed algorithm: Study Area 1, in which low-rise buildings are predominant; Study Area 2, in which relatively large gable-roof and hip-roof buildings are located; and Study Area 3, in which a mixture of high-rise and low-rise buildings coexist. Figures 6–8 present the building segmentation results of Study Areas 1, 2, and 3, respectively. Each image has an area of 1,000 × 1,000 pixels, which is equivalent to 250 m × 250 m. To examine segmentation performance, the commercial segmentation software, ENVI EX (Version 4.8) [32], was used for comparison. The software segments regions using gradient map and watershed algorithm [33]. The "feature extraction" function in this software 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. Figures 6–8 thus include the results generated using the proposed algorithm and those using ENVI EX.

Table 1. Experimental parameter values.

Process	Parameters	Value Used
Quantization and edge detection	Number of quantizations N_{disc}	3
	Quantization interval widths Δd_i (i = 1, ... , N_{disc})	40, 30, 20
	Number of offsets N_{off}	5
	Edge count T_{count1} and T_{count2}	5 and 3
	Minimum and maximum areas	50 and 30,000 pixels
	Minimum score for edge completion using filters shown in Figure 4, T_{count3} and T_{count4}	2 and 8
Segmentation and calculation of rectangular index	Minimum rectangular index	0.45
	Minimum and maximum distance between edges for rectangular index calculation, d_{edge_min}, d_{edge_max}	5 and 20 pixels
	Minimum valid length of rectangle	8 pixels
Selection regions according to rectangular index	Maximum ratio of overlapping area to original area for selecting areas overlapping with previously selected areas	0.2
	Minimum area for selecting areas overlapping with previously selected areas (same for Step (1))	50 pixels

To reduce the computation time, the labeling and edge detection was implemented in 50 × 50 pixel windows. These window images were extracted from each 1,000 × 1,000 pixel image as follows. First, the line and pixel positions of the upper left corner of the window were set to (0, 0), (0, 50), (0, 100), ..., (0, 950), (50, 0), ..., and (950, 950). After this, the positions were set to (25, 25), (25, 75), ..., (25, 925), (75, 25), ..., and (925, 925). The edges detected in all of the windows were then merged. Similarly, segmentation and the selection of regions were implemented in a 500 × 500 pixel window, and the results again merged. Finally, regions close to the boundaries of the window were put through the selection process a second time such that calculation errors ensuing from the merging of small-window results were almost completely negated.

Figure 5. (**a**) aerial photograph, (**b**) labeling from quantization using three different interval widths, (**c**) edge detection, (**d**) segmentation, and (**e**) final segmentation result. The left, central, and right images in (b), (c), and (d) were generated with Δd = 40, 30, and 20, respectively. A square over (a) indicates the area represented in (b).
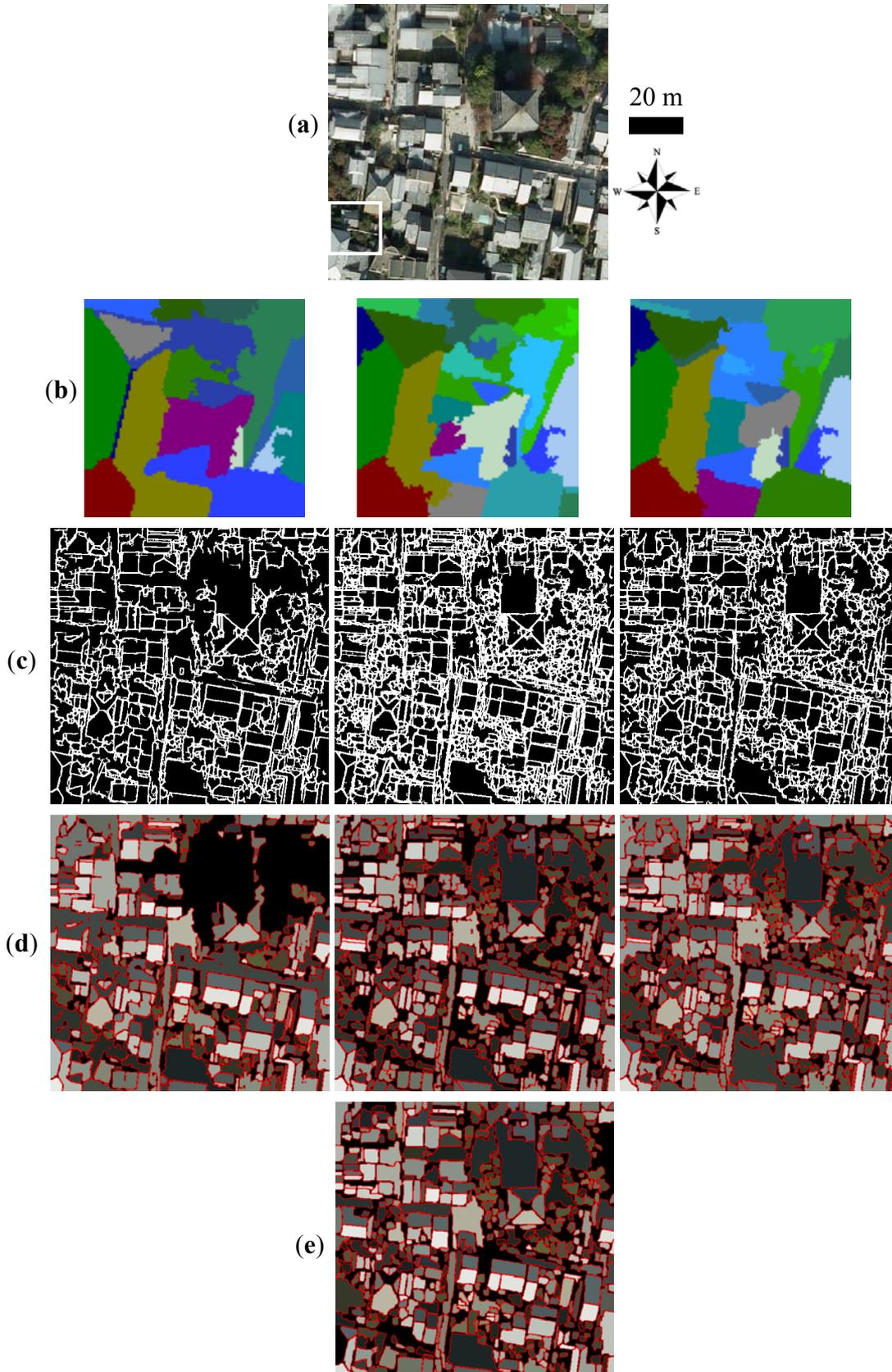
Figure 6. Comparison of building segmentation results in Study Area 1, in which low-rise buildings are predominant: (upper) aerial photograph, (lower left) segmentation result using the proposed algorithm, and (lower row) segmentation results using ENVI EX. Parameter values of "Scale Level" and "Merge Level" in ENVI EX were set to 50 and 80, respectively.

250 m

50 m

250 m
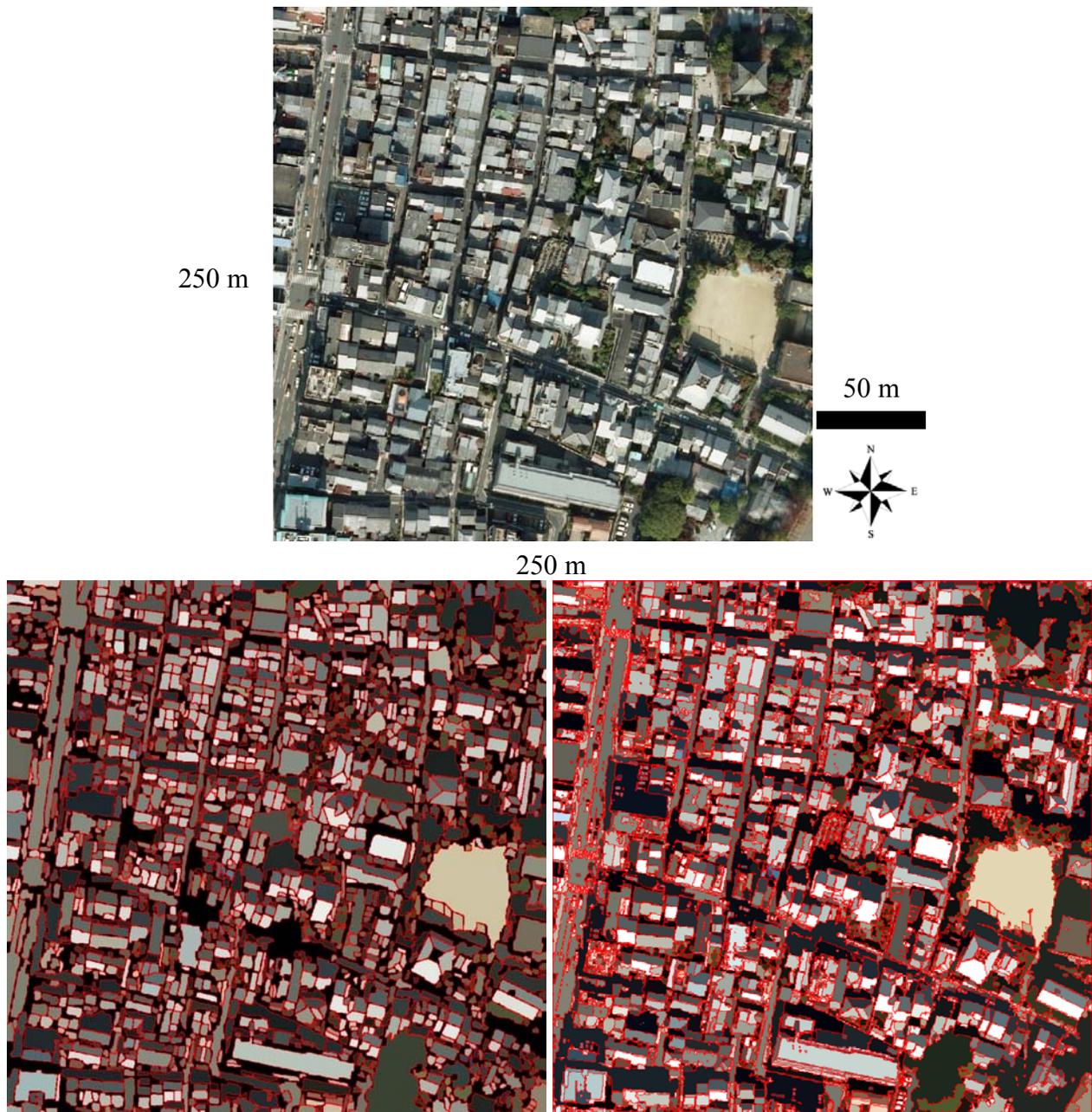
Figure 7. Comparison of building segmentation results in Study Area 2, in which relatively large gable-roof and hip-roof buildings are located. See Figure 6 for a description of each panel.
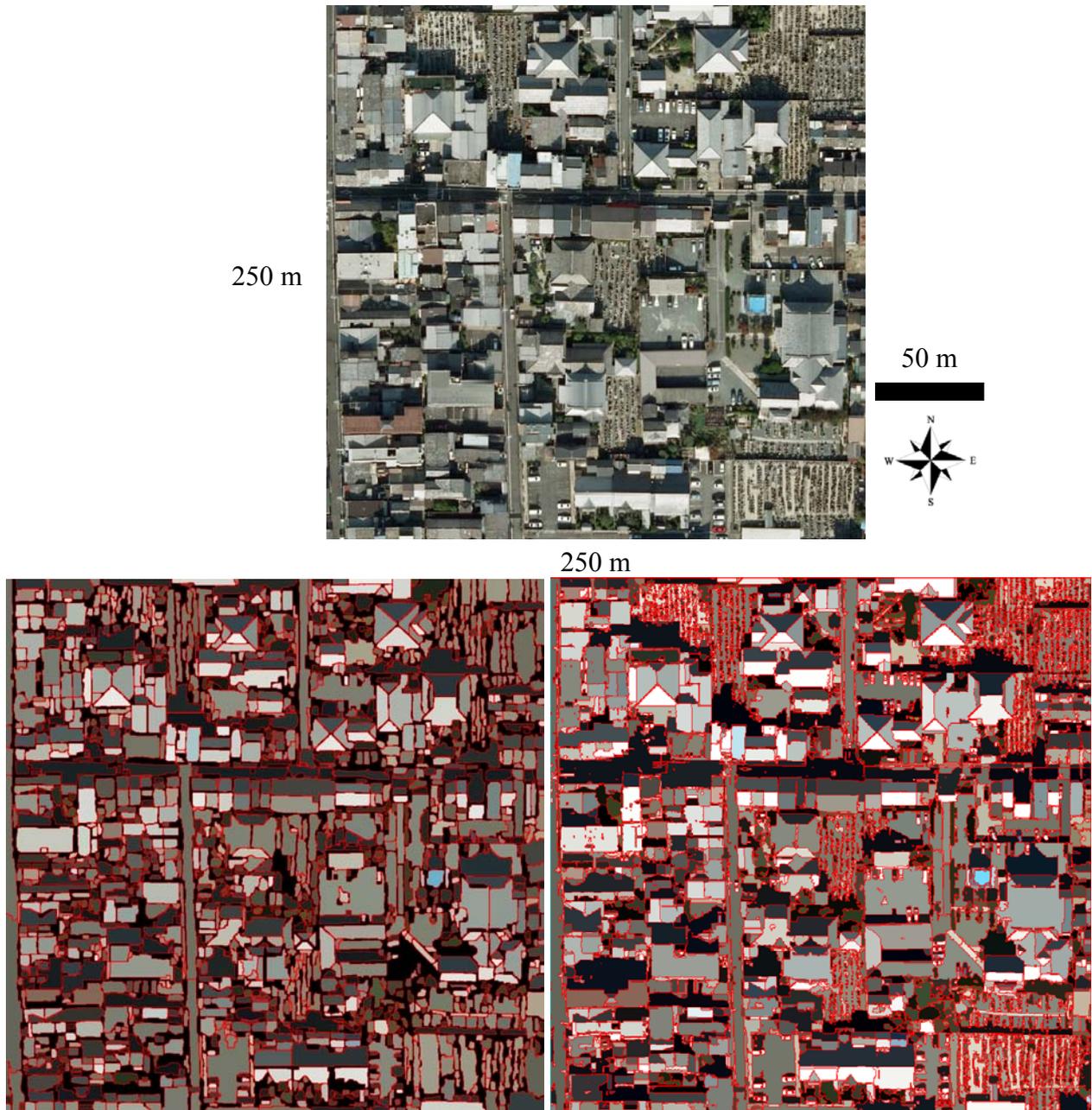
Figure 8. Comparison of building segmentation results in Study Area 3, in which a mixture of high- and low-rise buildings coexist. See Figure 6 for a description of each panel.
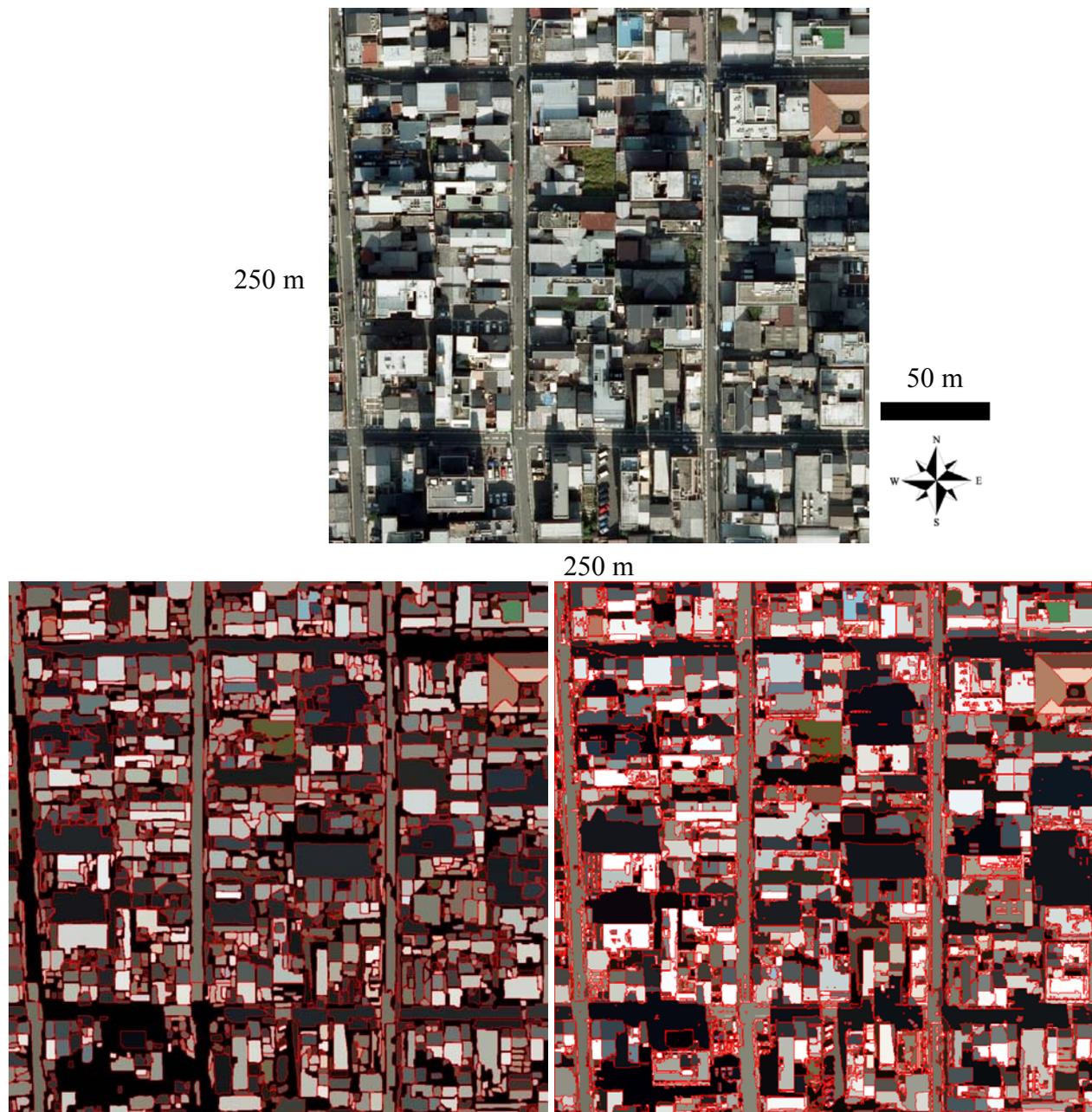
In the experiment, three interval widths, Δd_i = 40, 30, and 20, were selected for quantization. Although an attempt was made to complete unclear boundaries by using the filters in Figure 4, a large number of shadowed or roughly textured roofs were still not segmented correctly. Therefore, the edges detected using the three interval widths were merged and the filters in Figure 4 were then applied to complete the edges. Explicitly, three types of edges were used: edges detected with Δd_i = 40, edges detected with Δd_i = 20, and the combination of edges detected with Δd_i = 40, 30, and 20. The effect of this merging of results is discussed in Section 4.1.

Segmentation results were assessed in terms of shadowing and building types. Shadowing was split into three categories: unshadowed (less than or equal to 10% of the roof area was covered by shadow),

partially-shadowed (greater than 10%, but less than or equal to 50%, shadowing), and mostly-shadowed (greater than 50% shadowing) buildings. The buildings whose entire areas were included in Study Area 1 were classified into flat-, gable-, hip-, and slant-roof buildings. Reference buildings were manually identified.

Assessment was conducted on an entire-building basis. This meant that in the case of gable- and hip-roof buildings, assessment was independent of whether each roof was successfully segmented. Segmentation performance was also split into five categories: (1) a building is segmented from other buildings, and the error between the segmented and actual areas is within 10%; (2) a building is segmented from other buildings, and the error between the segmented and actual areas is greater than 10% and less than or equal to 50%; (3) a building is merged with one or more other buildings; (4) a building is merged with a road; and (5) the error between the segmented and actual area exceeds 50%. Therefore, a lower category number represents better segmentation performance.

Figure 9 shows the validation of the segmentation results for all buildings. Figures 10–12 then show the validation of the segmentation results for unshadowed, partially-shadowed, and mostly-shadowed buildings, respectively. Accuracy for validation was obtained by computing the ratio of total segmented area to area of reference building.
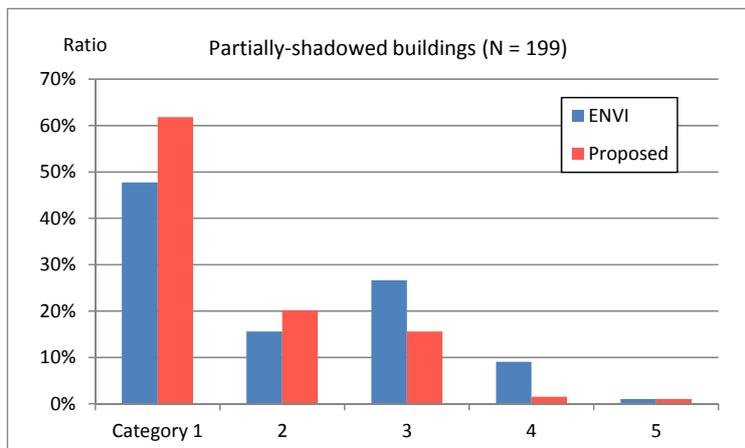
**Figure 9.** Verification of segmentation results for all buildings in Study Area 1.
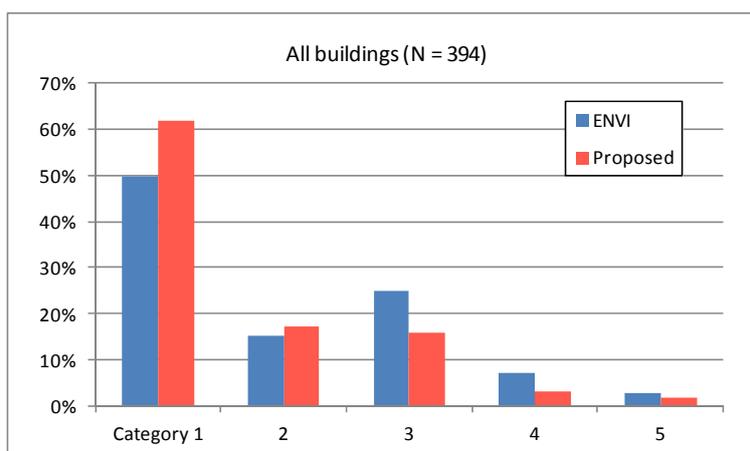
Figure 10. Verification of segmentation results for unshadowed (less than 10%) buildings in Study Area 1.
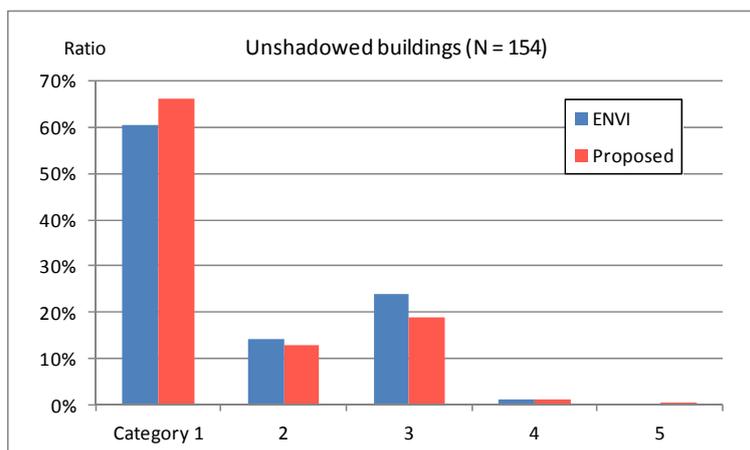
Figure 11. Verification of segmentation results for partially-shadowed (10% to 50%) buildings in Study Area 1.

Figure 12. Verification of segmentation results for mostly-shadowed buildings in Study Area 1.
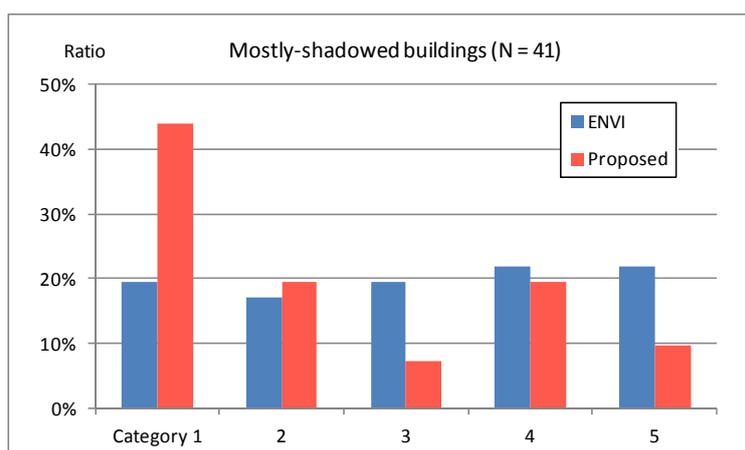
5. Discussions

5.1. Effect of Quantization and Edge Completion

The proposed algorithm merges regions segmented using the edges detected with different DN interval widths, Δd_i. 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. In the algorithm, regions with a high rectangular index are selected from the regions generated at each quantization. Figure 5 demonstrates that this selection procedure optimizes the local spatial scale for smoothing.

The proposed algorithm attempts to extract regions whose shape is close to being rectangular through the rectangular index calculated from a region's edges. However, for roughly textured roofs or in dense urban areas where building boundaries are often unclear, successful detection of complete edges is nontrivial. Failure to delineate boundaries in such circumstances reduces the accuracy of

building segmentation. As shown in Figure 13, the quantization of DNs and the combination of results for several interval widths in the proposed algorithm help to distinguish these roughly-textured roofs and unclear boundaries. Unsuccessful segmentation results, for example, where a building and road are merged, may have a lower rectangular index. Such results are excluded because the algorithm selects only those regions with a high rectangular index.

In addition to these factors, edge completion also contributes to the improvement of segmentation accuracy. Edges are completed by including those pixels that have high probability of being in an edge because they have neighboring edges pixels. Figure 14 shows that edge completion by using filters prevented shadowed roofs and buildings being merged with roads.

In ENVI EX, any of a number of edge operators can be used as a gradient operator [33]. To examine the edge detection performance, the Canny filter, which is a traditional powerful filter, was applied. The result of edges detected using the Canny is not included in this paper, but it was difficult to successfully extract edges of boundaries of partially-shadowed and mostly-shadowed buildings. In addition, the Canny extracted many edges from rough texture of roofs, which may lower the performance of the roof or building segmentation. Compared with this result, both quantization and edge completion of the proposed algorithm help in extracting more edges of building boundaries and less edges of roof texture.

Figure 15 shows the effect of another edge completion. As mentioned in Section 3, three types of edges were used in the algorithm: edges detected with $\Delta d_i = 40$, edges detected with $\Delta d_i = 20$, and the combination of edges detected with $\Delta d_i = 40, 30,$ and 20. Figure 15(d) shows that the combined edges are effective to segment shadowed buildings. However, segmentation using the combined edges was found to typically extract smaller regions compared with segmentation using a single interval width. Therefore, selection among regions segmented using edges found with both single and combined interval widths can generate reasonable results.

Consequentially, as shown in Figures 9 to 12, the proposed algorithm produces higher accuracy segmentation than the existing algorithm. In particular, in the cases of partially- and mostly-shadowed buildings, the proposed algorithm performs much better than the existing algorithm. As shown in Figures 11 and 12, in cases of partially-shadowed (10% to 50%) buildings and mostly-shadowed buildings, the ratios of Category 1 (the error between the segmented and actual areas was within 10%) obtained by using the proposed algorithm were 12% and 24% higher than the ones obtained by using ENVI EX, respectively. Low gable-roof buildings in a dense urban area have a high likelihood of being partially- or mostly-shadowed. However, it has been demonstrated that the proposed algorithm can accurately segment highly shadowed buildings.

Figure 13. Comparison of building segmentation results: (**left**) aerial photograph, (**middle**) segmentation results using the proposed algorithm, and (**right**) segmentation results using ENVI EX. (**a**) and (**b**) Results for Study Area 1, (**c**) results for Study Area 2, and (**d**) results for Study Area 3.
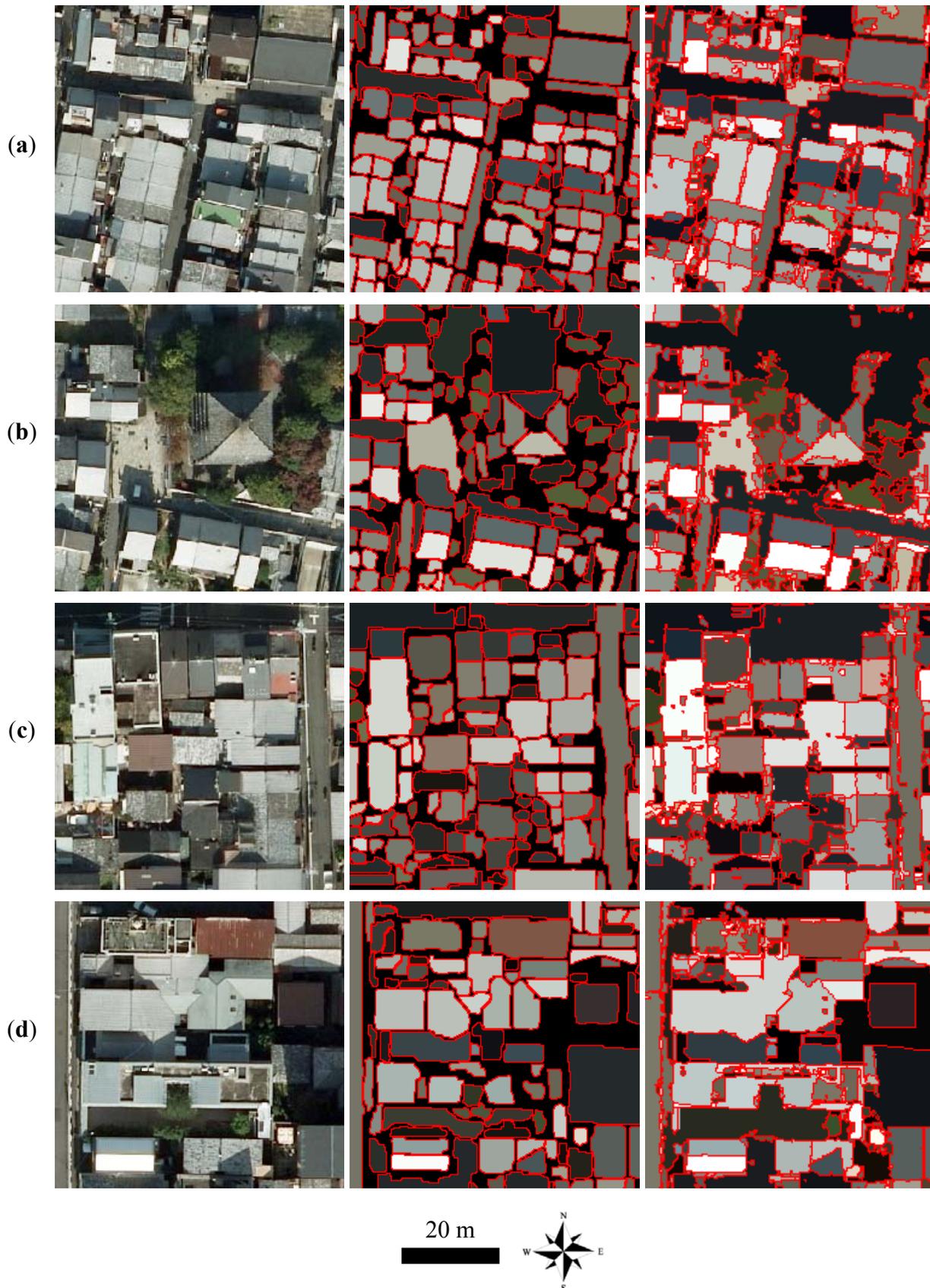
Figure 14. Edge completion using filters: (**a**) non-completed edges, (**b**) segmentation result using non-completed edges, (**c**) completed edges, and (**d**) segmentation result using completed edges. All results were generated with $\Delta d_i = 40$.

(**a**)

(**b**)

(**c**)

(**d**)

20 m
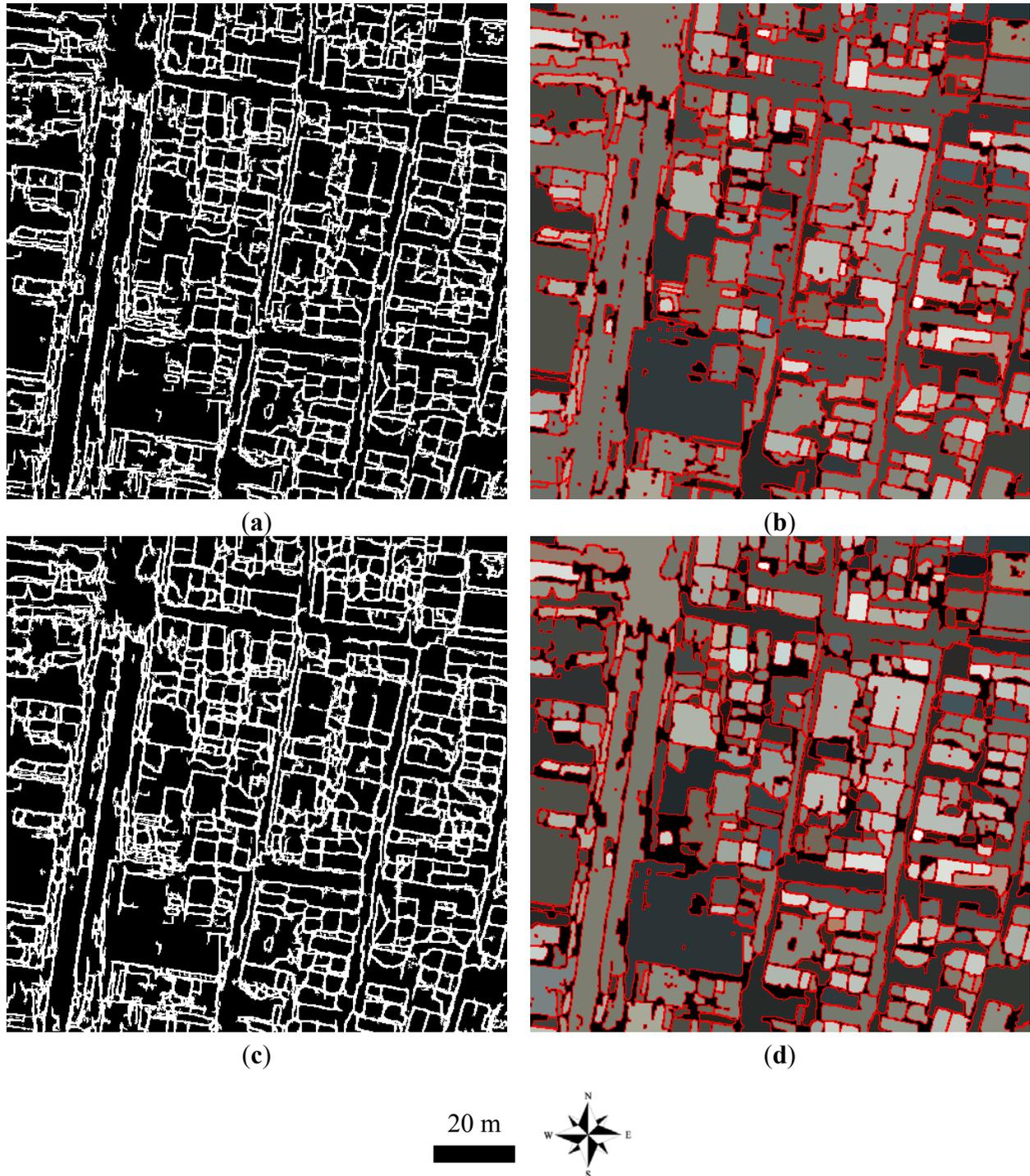
Figure 15. Edge completion by using combined edges: (**left**) extracted edges and (**right**) segmentation results using these edges. (**a**) Results with $\Delta d_i = 40$, (**b**) results with $\Delta d_i = 30$, (**c**) results with $\Delta d_i = 20$, and (**d**) combined edges when edges with $\Delta d_i = 40$, 30, and 20 were merged.
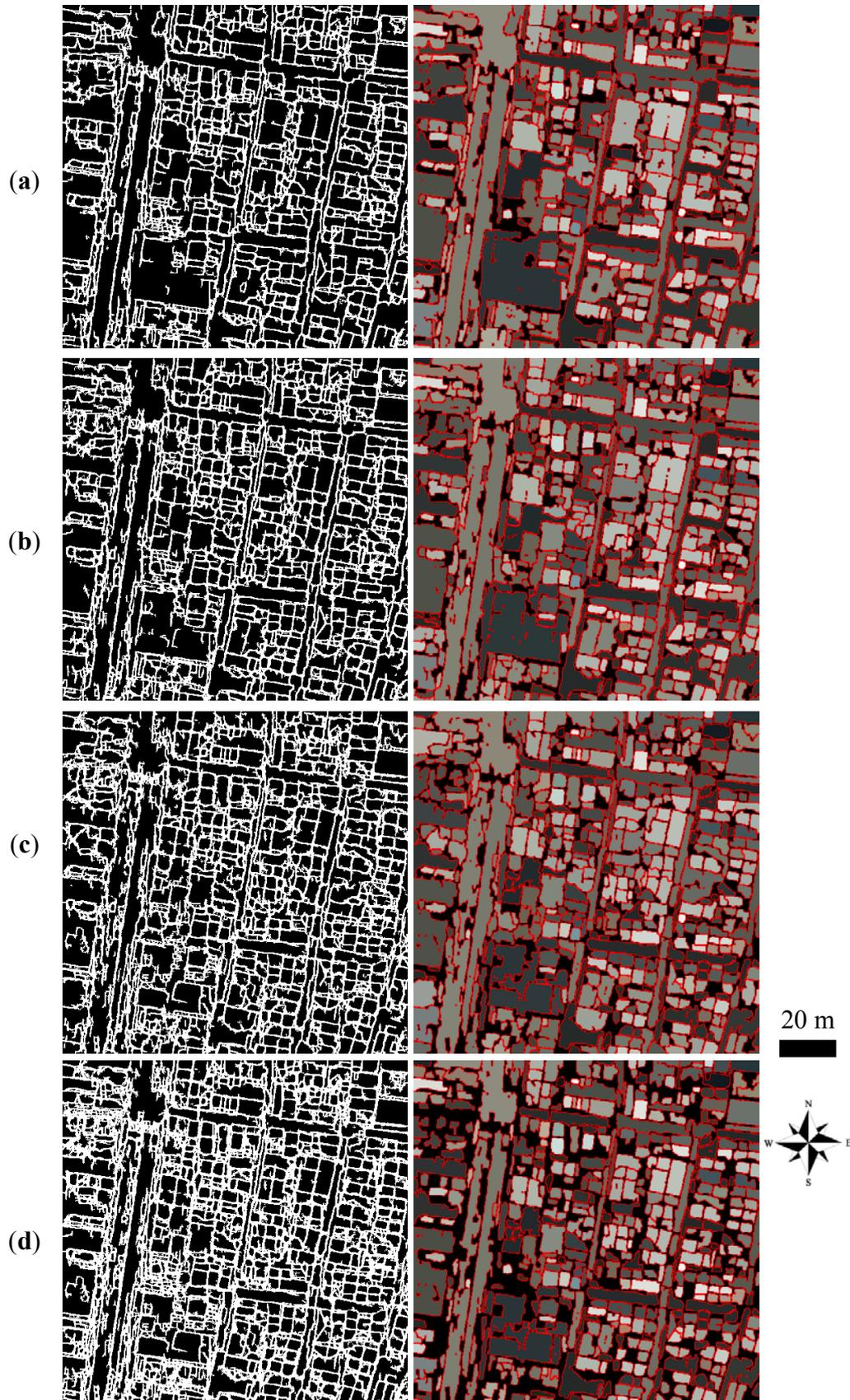
5.2. Rectangular Index

The rectangular index selects an optimal region at a specific location from a number of candidates. Because the author's focus in this research was on urban areas, this index is considered appropriate for extracting buildings. In spite of this, as shown in Figures 13(b), triangular regions of hip-roofs are also extracted by using the proposed algorithm. A perfect triangle's rectangular index is only 0.5, and so the proposed algorithm does not prioritize triangular regions for selection. The reason for this successful extraction may be that neighboring regions were already successfully extracted. Under certain interval widths and offsets, a triangular region and its neighbor are often merged. However, under different values of these parameters, they may become separated. A merged region that includes a triangular region might not be selected, however, since they tend to have a lower rectangular index. Instead, the regions around the triangular region are extracted, and extraction of the triangular region then follows this. Whether the proposed algorithm is also able to segment circular roofs or buildings could not be confirmed, because none were found in the study areas. However, based on the successful segmentation of triangular regions, it may be possible to correctly segment such roofs and buildings.

However, selection based on the rectangular index presents a problem. The proposed algorithm extracts regions according to the rectangular index without considering a region's area. As a result, small regions with a high rectangular index are selected above large regions with a lower rectangular index, even though the large region may be more suitable for delineating the building. An approach to prioritize such large regions by applying a correction to the rectangular index was therefore examined. As a result, a greater number of large regions corresponding to roads or vegetation and a lesser number of building regions were selected. The idea of a correction to the rectangular index may be useful for certain purposes (e.g., segmentation of a number of buildings on a district level). However, issues remain that require consideration: for example, an appropriate functional form for the correction and adjustment of the coefficients in such a function. Hence, the results shown in the present paper were generated without this correction.

Discussion now turns to calculation of the rectangular index. When PCA was employed in rectangular index calculations, many slate roofs were divided into small regions or parts of the slate roofs were missed. In contrast, over-merged roofs were also found. Application of PCA in rectangular index calculations generated axes that were far from being parallel to the rectangular sides. Therefore, because segmentation results using PCA were found to be unstable, the main and sub axes used in rectangular index calculations were determined by the procedure explained in Section 3.1. Although the thresholds must be optimized empirically, this approach was found to have higher stability than PCA.

5.3. Optimization of Parameters

Among the parameters listed in Table 1, final segmentation results are sensitive to those related to quantization, edge detection, and completion. In particular, DN interval widths were repeatedly examined during the experiment. These intervals are dependent on the brightness contrast, and empirical determination of the intervals through a number of investigations may be necessary. The design of filters for edge completion is dependent on the objects to be segmented. In the experiment,

filters were selected to complete linearly-aligned edges because rectangular buildings were dominant in the study areas. In the case of extraction of round buildings, the filters should be designed to complete curved edges.

5.4. Computation Time

Labeling of regions after quantization requires computation time. Therefore, the technique described in Section 3 of splitting the area into small windows during edge detection and segmentation was adopted. Comparing this segmentation with the result without such a split, no significant difference was found. The computation time for different-sized images is shown in Figure 16. This experiment was conducted using a PC with an Intel Core i7 (3.20 GHz) processor and 6 GB memory. The computation time is almost proportional to the number of pixels, and the proposed algorithm is shown to be useful.

Figure 16. Computation time for different-sized aerial images.
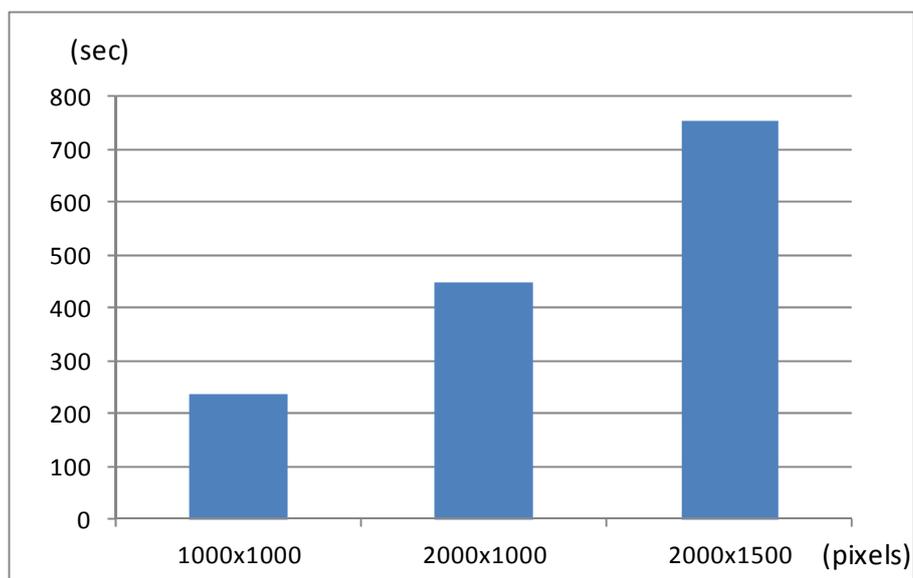
5.5. Applications

The author's interest in conducting this research is to generate 3D building models using airborne LiDAR data and the segmented results obtained using the proposed algorithm. The author developed a 3D building modeling algorithm that uses the results of building segmentation from aerial photographs. With the information of roofs and buildings, the accuracy of 3D building models was improved even in the dense urban areas where houses that have slant roofs are located close to each other, and their heights are similar [34]. In addition, the proposed algorithm can be applied to the generation of 2D maps of buildings. Such 2D building maps are useful for applications that require rapid map generation to ascertain the status of an urban area without the need for high accuracy. For example, assessment of damage caused by a natural disaster—an earthquake, flood, or tsunami—is a conventional application. In assessing the damage caused by the Great East Japan Earthquake on 11 March 2011, 2D thematic maps were useful for national and local governments. However, the majority of these maps were generated through manual interpretation. Compared with existing

algorithms, segmented results by the proposed algorithm are less affected by shadows, and thus manual correction of the results is greatly reduced.

Ideal processing of 2D building maps should automatically exclude vegetation, whereas vegetation was not removed in this research. The timing of vegetation removal was a complex issue. Removal in the preprocessing stage of pixels whose DNs are similar to those of vegetation was examined. However, this approach removed the vegetation pixels covering buildings and roads. As a result, regions considerably smaller than the actual buildings were extracted, or regions were not extracted because their areas were below the threshold. Another examined approach was to retain vegetation pixels during segmentation and remove regions having a high probability of being vegetation at the end. This approach was successful, while some large vegetated regions were not removed. However, the removal of red vegetation while maintaining red roofs was still difficult. Because vegetation removal is a key factor in various applications of the proposed algorithm, it will be examined in the near future.

6. Conclusions

In this paper, an algorithm to segment buildings, including shadowed buildings, from aerial photographs of dense urban areas was proposed. To distinguish roofs having a rough texture, DNs are quantized into a number of quantum values. Quantization using several interval widths is applied during segmentation, and for each quantization, areas with homogeneous values are labeled in an image. Edges determined from the homogeneous areas obtained at each quantization are merged, and frequently observed edges are extracted. By using a rectangular index, regions whose shapes are close to being rectangular are selected as buildings. Finally, pixels that have the potential to be part of an edge from the context of neighboring pixels are added to edges in order to improve segmentation accuracy. Quantization using three interval widths was applied in the experiment, and the main factors leading to successful segmentation of shadowed roofs were (1) the combination of different quantization results, (2) selection of buildings according to the rectangular index, and (3) edge completion. Crucially, owing to these three factors, the scale of the spatial filtering is optimized with respect to the size of building roofs in a locality. In addition, even though the proposed algorithm does not prioritize triangular regions, such regions are extracted. Owing to selection based on the rectangular index, the regions around a triangular region were extracted, and as a result, the triangular regions were also extracted. The experimental results showed that the proposed algorithm generated better segmentation results than an existing algorithm. In particular, in the cases of partially-shadowed (10% to 50%) buildings and mostly-shadowed buildings, the ratios of category that the error between the segmented and actual areas was within 10% obtained by using the proposed algorithm were 12% and 24% higher than the ones obtained by using ENVI EX, respectively. Therefore, the proposed algorithm is considered to be useful for conducting building segmentation for various purposes. Although the computation time for segmentation was deemed reasonable, this should be greatly reduced through future investigation.

Acknowledgments

References

1. Elberink, S.O. Target Graph Matching for Building Reconstruction. In *Proceedings of Laserscanning'09*, Paris, France, 1–2 September 2009; In *International Archives of the Photogramm., Remote Sensing and Spatial Information Sciences*; 2009; Volume XXXVIII, 3/W8, pp. 49-54.

2. Steuer, H. Height Snakes: 3D Building Reconstruction from Aerial Image and Laser Scanner Data. In *Proceedings of 2011 Urban Remote Sensing Joint Event*, Munich, Germany, 11–13 April 2011; pp. 113-116.

3. Huber, M.; Schickler, W.; Hinz, S.; Baumgartner, A. Fusion of LIDAR Data and Aerial Imagery for Automatic Reconstruction of Building Surfaces. In *Proceedings of 2nd GRSS/ISPRS Joint Workshop on "Data Fusion and Remote Sensing over Urban Areas"*, Berlin, Germany, 22–23 May 2003; pp. 82-86.

4. Sohn, G.; Dowman, I. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS J. Photogramm.* **2007**, *62*, 43-63.

5. Rottensteiner, F.; Trinder, J.; Clode, S.; Kubik, K. Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis. *ISPRS J. Photogramm.* **2007**, *62*, 135-149.

6. Awrangjeb, M.; Ravanbakhsh, M.; Fraser, C.S. Automatic detection of residential buildings using LIDAR data and multispectral imagery. *ISPRS J. Photogramm.* **2010**, *65*, 457-467.

7. Tso, B.; Mather, P.M. *Classification Methods for Remotely Sensed Data*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2009.

8. Soille, P.; Pesaresi, M. Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2042-2055.

9. Benediktsson, J.A.; Pesaresi, M.; Amason, K. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1940-1949.

10. Benediktsson, J.A.; Palmason, J.A.; Sveinsson, J.R. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 480-491.

11. Bellens, R.; Gautama, S.; Martinez-Fonte, L.; Philips, W.; Chan, J.C.-W.; Canters, F. Improved classification of VHR images of urban areas using directional morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2803-2813.

12. Tuia, D.; Pacifici, F.; Kanevski, M.; Emery, W.J. Classification of Very High Spatial Resolution imagery using mathematical morphology and Support Vector Machines. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3866-3879.

13. Novack, T.; Esch, T.; Kux, H.; Stilla, U. Machine learning comparison between WorldView-2 and QuickBird-2-simulated imagery regarding object-based urban land cover classification. *Remote Sens.* **2011**, *3*, 2263-2282.

14. Moskal, L.M.; Styers, D.M.; Halabisky, M. Monitoring urban tree cover using object-based image analysis and public domain remotely sensed data. *Remote Sens.* **2011**, *3*, 2243-2262.

15. Tsai, Y.H.; Stow, D.; Weeks, J. Comparison of object-based image analysis approaches to mapping new buildings in Accra, Ghana using multi-temporal QuickBird satellite imagery. *Remote Sens.* **2011**, *3*, 2707-2726.

16. Canny, J. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *PAMI-8*, 679-698.

17. Wan, T.; Canagarajah, N.; Achim, A. Segmentation of noisy colour images using cauchy distribution in the complex wavelet domain. *IET Image Process.* **2011**, *5*, 159-170.

18. Zhong, J.; Ning, R. Image denoising based on wavelets and multifractals for singularity detection. *IEEE Trans. Image Process.* **2005**, *14*, 1435-1447.

19. Rezaee, M.R.; van der Zwet, P.M.J.; Lelieveldt, B.P.E.; van der Geest, R.J.; Reiber, J.H.C. A multiresolution image segmentation technique based on pyramidal segmentation and fuzzy clustering. *IEEE Trans. Image Proces.* **2000**, *9*, 1238-1248.

20. Cheng, H.; Bouman, C.A. Mutliscale Bayesian segmentation using a trainable context model. *IEEE Trans. Image Processing* **2001**, *10*, 511-525.

21. Esch, T.; Thiel, M.; Bock, M.; Roth, A.; Dech, S. Improvement of image segmentation accuracy based on multiscale optimization procedure. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 463-467.

22. Berlemont, S.; Olivo-Marin, J.-C. Combining local filtering and multiscale analysis for edge, ridge, and curvilinear objects detection. *IEEE Trans. Image Process.* **2010**, *19*, 74-84.

23. Bouman, C.A.; Shapiro, M. A mutliscale random field model for Bayesian image segmentation. *IEEE Trans. Image Process.* **1994**, *3*, 162-177.

24. Paget, R.; Longstaff, I.D. Texture synthesis via a noncausal nonparametric multiscale Markov random field. *IEEE Trans. Image Process.* **1998**, *7*, 925-931.

25. Liang, K.-H.; Tjahjadi, T. Adaptive scale fixing for multiscale texture segmentation. *IEEE Trans. Image Process.* **2006**, *15*, 249-256.

26. Tonazzini, A.; Bedini, L.; Salerno, E. A Markov model for blind image separation by a mean-field EM algorithm. *IEEE Trans. Image Process.* **2006**, *15*, 473-482.

27. Ding, J.; Ma, R.; Chen, S. A scale-based connected coherence tree algorithm for image segmentation. *IEEE Trans. Image Process.* **2008**, *17*, 204-216.

28. Chien, S.Y.; Ma, S.Y.; Chen, L.G. Efficient moving object segmentation algorithm using background registration technique. *IEEE T. Circ. Syst. Vid.* **2002**, *12*, 577-586.

29. Tsai, V.J.D. A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1661-1671.

30. Ma, H.; Qin, Q.; Shen, X. Shadow Segmentation and Compensation in High Resolution Satellite Images. In *Proceedings of 2008 IEEE International Geosci. Remote Sensing Symposium*, Boston, MA, USA, 7–11 July 2008; pp. II-1036-II-1039.

31. Deng, Y.; Manjunath, B.S. Unsupervised segmentation of color-texture regions in images and video. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 800-810.

32. 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 (accessed on 5 June 2011).

33. Jin, X. Segmentation-based image processing system. US Patent 20,090,123,070, filed 14 November 2007, and issued 14 May 2009.

34. Susaki, J. Fusion of airborne LiDAR data and aerial photographs for automatic modeling of buildings in dense urban areas. *IEEE Trans. Geosci. Remote Sens*. **2012**, in preparation.