AN IMPROVED SNAKE MODEL FOR REFINEMENT OF LIDAR-DERIVED BUILDING ROOF CONTOURS USING AERIAL IMAGES

Qi Chen [a], Shugen Wang [b], Xiuguo Liu [a, *]

[a] Faculty of Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, P. R. China
- daoqiqi@gmail.com, liuxg318@163.com
[b] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430049, P. R. China - wangsg@whu.edu.cn

Commission III, WG III/4

KEY WORDS: Building Roof Contour, Refinement, Aerial Images, Snake Model, Greedy Algorithm

ABSTRACT:

Building roof contours are considered as very important geometric data, which have been widely applied in many fields, including but not limited to urban planning, land investigation, change detection and military reconnaissance. Currently, the demand on building contours at a finer scale (especially in urban areas) has been raised in a growing number of studies such as urban environment quality assessment, urban sprawl monitoring and urban air pollution modelling. LiDAR is known as an effective means of acquiring 3D roof points with high elevation accuracy. However, the precision of the building contour obtained from LiDAR data is restricted by its relatively low scanning resolution. With the use of the texture information from high-resolution imagery, the precision can be improved. In this study, an improved snake model is proposed to refine the initial building contours extracted from LiDAR. First, an improved snake model is constructed with the constraints of the deviation angle, image gradient, and area. Then, the nodes of the contour are moved in a certain range to find the best optimized result using greedy algorithm. Considering both precision and efficiency, the candidate shift positions of the contour nodes are constrained, and the searching strategy for the candidate nodes is explicitly designed. The experiments on three datasets indicate that the proposed method for building contour refinement is effective and feasible. The average quality index is improved from 91.66% to 93.34%. The statistics of the evaluation results for every single building demonstrated that 77.0% of the total number of contours is updated with higher quality index.

1. INTRODUCTION

Buildings are the most significant component of the urban scence. The location information of buildings is widely applied in many fields, such as urban planning, land investigation, change detection and military reconnaissance. As an effective means to acquire high-resolution elevation data, airborne light detection and ranging (LiDAR) is increasingly frequently used in building detection. Considering that this technique has an obvious advantage on height precision，building detection based on LiDAR usually has a higher degree of automation, which makes these types of detection methods able to perform well in many applications (Dorninger and Pfeifer, 2008; Awrangjeb and Fraser, 2014; Mongus et al., 2014; Niemeyer et al., 2014).

Currently, the demand on building contours at a finer scale (especially in urban areas) has been raised in a growing number of studies such as urban environment quality assessment, urban sprawl monitoring and urban air pollution modelling. However, the quality of the building contours extracted from LiDAR are always restricted by the scan resolution of the point cloud. In spite of the design level of the airborne LiDAR equipment is continuously improved in recent years, the ground resolution of the point cloud still falls behind the aerial images when acquiring data at a similar height. This difference means that even if the building detection results are completely correct in LiDAR, there is still some room for the accuracy improvement of the contours, which makes the refinement of LiDAR-derived building contours a problem of concern.

Using the edge information from images with higher resolution is a feasible way to perform the building contour refinement. The refinement based on image straight-line features is one of the traditional methods (Cheng et al., 2011; Awrangjeb et al., 2013; Li et al., 2013; Dal Poz, 2013). By updating the initial building contours with the building edge line segments extracted from high-resolution images, more accurate contours can be obtained through these methods. However, these methods are generally too dependent on the results of line feature extraction, the false features and difficult-extracted features can both pose challenges for the refinement process.

In general, the LiDAR-derived building contours should be simplified or regularized before being refined. Some assumptions including rectangularity criterion (Galvanin and Dal Poz, 2012) and principal orientation constraint (Cheng et al., 2011; Zhao et al., 2016) are frequently used for the contour simplification and regularization. However, since the increasing complexity of buildings in various applications, these assumptions may not be completely suitable for many buildings (as shown in Figure 1). Refining the building contour with an inappropriate constraint would probably lead to erroneous results.

Figure 1. Building contours that contradict against the commonly-used regularization assumptions

---

* Corresponding author

Compared with the abovementioned building contour processing methods, the advantage of the Snake model lies in the ability of integrating the input data, initial estimation, optimal contour generation and knowledge-based constraint into a single process. It makes this model applied widely in building contour extraction and optimization. The Snake model is more often employed for building contour extraction with high-resolution images as a single data source (Peng et al., 2005; Li et al., 2012; Fazan and Dal Poz 2013). However, considering that this model largely relies on a good initial contour, refined contour with high quality may be difficult to obtain when only using images. Kabolizade et al. (2010) proposed an improved Snake model to refine the LiDAR-derived building contours with aerial images. In their approach, better results were achieved with the improved model than the traditional Snake model. But it should be noted that they only used single aerial image to perform the refinement, and the test objects are mainly buildings with rectangular corners.

In this paper, we modified the classic Snake model to make it adaptable to more complicated building contours extracted from LiDAR. Assuming that high-precision building detection results have been obtained from LiDAR point cloud, the target of this study is to further improve the quality of building contours using multi-view aerial images. Our research will partly reflect the potential of automatically generating high-quality building contour based on combined utilization of airborne LiDAR and aerial imagery.

The remainder of this paper is organized as follows. Section 2 describes the proposed method. The experimental and evaluation results are presented and discussed in Section 3. Conclusions are given in Section 5.

2. METHODOLOGY

2.1 Traditional Snake Model

The basic principle of Snake model is to find an optimal contour that satisfies the minimization constraint of internal and external energy, where the internal energy ensures the contour to be smooth and continuous, and the external energy drives the contour toward the edge of the target object. In the actual digital image processing, the contour in Snake model is discretized into a series of control nodes $(v_1, v_2, v_3, \cdots, v_n)$, the energy function of the contour is expressed as the sum of the internal energy E_{int} and the external energy E_{ext} of each node:

$$E_{snake} = \sum_{i=1}^{n}(E_{int}(i) + E_{ext}(i)) \qquad (1)$$

where E_{int} is composed of the elastic energy $E_{elastic}$ and the bending energy $E_{bending}$, which can be expressed as:

$$\begin{aligned}
E_{elastic}(i) &= \alpha(|v_i - v_{i-1}|)^2/2 \\
E_{bending}(i) &= \beta(|v_{i-1} - 2v + v_{i+1}|)^2/2 \\
E_{int}(i) &= E_{elastic}(i) + E_{bending}(i)
\end{aligned} \qquad (2)$$

where α and β are weight values for $E_{elastic}$ and $E_{bending}$, respectively. The elastic energy aims to prevent the contour from being stretched, while the bending energy gives the contour a "restitution" force when bent or twisted, driving it to be a smooth curve or a straight line.

2.2 Improved Snake Model

Since most of the buildings in reality have distinct corners, which actually contradicts the smooth and continuous constraint of the Snake model. In order to better meet the requirements of building contour refinement, some improvements are made for the traditional Snake model in our approach.

2.2.1 Internal Energy Term: As shown in Figure 2(a), a number of redundant nodes often exist on the initial building contour derived from LiDAR point cloud. Line simplification methods such as Douglas-Peucker algorithm (Douglas and Peucker, 1973) are generally adopted to reduce the redundant nodes. But for complex buildings that have many detailed edge information on the roof, a perfect threshold may not exist for the simplification: a large threshold leads to loss of the original details, while a small value would have little practical effect of simplification.
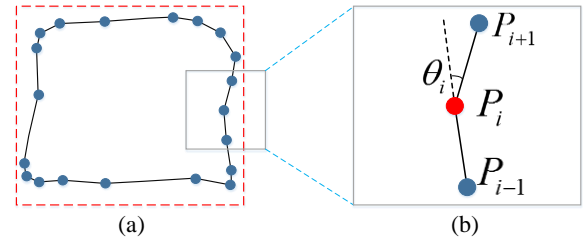

Figure 2. Determination of the internal energy item. (a) The redundant nodes of the initial LiDAR-derived building contour. (b) Definition of the deviation angle.

In order to preserve the details of the initial LiDAR-derived contour as much as possible, a relatively small threshold (0.5 m) is used to perform the contour simplification. Meanwhile, in this approach, the contour will be simplified in each iteration during the contour optimization process until the optimal result is obtained. To constrain the redundant nodes to a reasonable position, a measurement named deviation angle θ is introduced to redefine the internal energy term. Figure 2(b) shows three consecutive nodes of a contour, which are $P_{i-1}(x_{i-1}, y_{i-1}, z_{i-1})$, $P_i(x_i, y_i, z_i)$ and $P_{i+1}(x_{i+1}, y_{i+1}, z_{i+1})$. The deviation angle $θ_i$ can be calculated using the two vectors formed by these three nodes. Considering that the elevation difference of the roof may easily create an obvious angle between the adjacent nodes, only the 2D coordinates are used for calculating the deviation angle:

$$\begin{aligned}
\vec{V_i} &= (x_i - x_{i-1}, y_i - y_{i-1}) \\
\vec{V_{i+1}} &= (x_{i+1} - x_i, y_{i+1} - y_i) \\
\theta_i &= \arccos\left(\frac{\vec{V_i} \cdot \vec{V_{i+1}}}{|\vec{V_i}||\vec{V_{i+1}}|}\right)
\end{aligned} \qquad (3)$$

where $\vec{V_i}$ and $\vec{V_{i+1}}$ are the 2D vectors calculated.

The irregularity of the initial contour is mainly caused by the redundant nodes, the reason that these redundant nodes cannot be simplified is the existence of the abovementioned deviation angle. When the contour becomes more regular, the accumulated value of the deviation angles will tend to decrease gradually and converge to a stable value eventually. Thus, the internal energy term of the building contour can be represented as follows:

$$E_{int} = \sum_{i=1}^{n} \theta_i \qquad (4)$$

2.2.2 External Energy Term: The internal energy term only provides a constraint that makes the contour become regular when being moved, while the definition of the external energy term aims at improving the contour accuracy with the use of external data such as aerial imagery. In our approach, the external energy term includes gradient energy $E_{gradient}$ and area energy E_{area}:

$$E_{ext} = E_{gradient} + E_{area} \qquad (5)$$

In general, the building contour accuracy can be improved based on the image edge information. The image gradient is the primary form of the edge information, based on which edge-point or straight-line features can be extracted. However, compared with point or line features, the continuity of gradient information makes it much easier to be applied in the Snake model.

Since the aerial images may have large overlapping area, a same building would probably appear on multiple images. Due to the different imaging angle, building edges that cannot be completely captured from a single image, may be largely complemented by the overlapped images. Figure 5 shows the original and gradient images of a building on two overlapped images. It can be seen that there is a significant complementary effect of edge information between the two images. Therefore, multi-view images will be adopted for the contour refinement in this study.
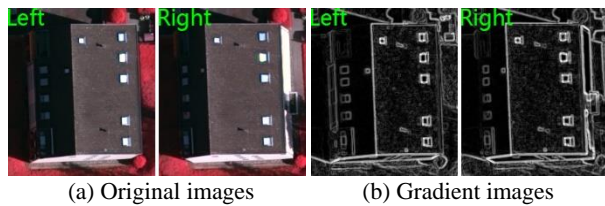

(a) Original images   (b) Gradient images

Figure 3. The complementary effect of edge information between the overlapped images

In order to calculate the gradient energy, the 3D contour should be projected onto the images using collinearity equation, and then the line segments of the contour should be rasterized. In our approach, the rasterization is performed using Bresenham algorithm (Bresenham, 1977), and the image gradient is extracted with Sobel operator. Gaussian blurring is performed to reduce the image noise before gradient calculation. The gradient energy will be calculated with the following formulas:

$$E_{gradient} = -\frac{1}{N \cdot C} \sum_{j=1}^{N} E_{gradient}(j)$$
$$E_{gradient}(j) = \sum_{k=1}^{n} |\nabla I_k(x, y)|^2 \qquad (6)$$

where $E_{gradient}(j)$ is the total gradient energy of the contour calculated from the jth image, N is the number of images used for gradient extraction, C is the horizontal projected perimeter of the 3D contour. $\nabla I_k(x, y)$ is the gradient value of the kth pixel of the rasterized 2D contour on the image, n is the number of pixels in the rasterized contour.

Owing to the existence of detailed structures such as skylights and chimneys on the roof, there are also some pixels inside the building contour having strong gradient. As a result, the contour may still move toward these kind of pixels when only using gradient as the external constraint. Using the fact that the area of

the LiDAR-derived initial contour is smaller than that of the actual contour in most cases, in our approach, an approximate method is proposed to estimate the area difference between the initial contour and the actual contour. As shown in Figure 4, the red lines represent the LiDAR derived contour, the green lines represent the actual contour, ΔGSD is the difference between the average scanning resolution of LiDAR and the image resolution. Assuming that the LiDAR points collected on the roof are symmetrically distributed, and considering that ΔGSD is a rather small value compared with the roof edge, the shadow area can be ignored. Therefore, the area difference between the initial contour and the actual contour can be estimated by multiplying ΔGSD and the perimeter of the initial contour.
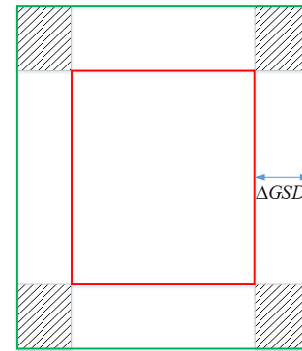

Figure 4. Area difference estimation between the LiDAR-derived contour and the actual contour

Base on the abovementioned assumptions, it can be considered that the area increment of the contour during the refinement process should be close to the estimated area difference. Thus, the area energy term is defined as follows:

$$\hat{A} = A_0 + \Delta GSD \cdot C$$
$$E_{area} = \frac{|A - \hat{A}|}{\Delta GSD \cdot C} \qquad (7)$$

where A_0 is the horizontal projected area of the LiDAR-derived 3D contour, $\hat{A}$ is the estimated area of the actual contour, A is the area of the current contour being refined. It can be noted that larger difference between the current contour area and the estimated area generates a greater area energy.

To sum up, the total energy E_{snake} of a building contour can be represented as:

$$E_{snake} = \alpha E_{int} + \beta E_{gradient} + \gamma E_{area} \qquad (8)$$

where α, β and γ are weights assigned to internal energy, gradient energy and area energy, respectively. In our approach, these values are set to 1, 10 and 1, respectively.

2.2.3 Convergence Criterion: Since the input data in this study are 3D vector contours rather than 2D rasterized contours, the process flow of building contour refinement is different from many other related studies based on Snake model. The key processes are as follows:

1) The initial building contours are extracted from the classification results of LiDAR point cloud.
2) The internal energy and area energy of the initial contour are calculated according to their definition, the gradient energy is calculated by projecting the contour onto multi-view images, and then the total energy can be calculated.

3) The nodes of the 3D contour are moved, and the shapes of the corresponding 2D contours on images are updated simultaneously. The total energy of the current contour is calculated.

4) Step 3) is repeated, and by minimizing the total energy of the contour, the optimal refined contour is generated.

The convergent iteration of the Snake model can be considered as a process of seeking the best solution based on dynamic optimization. In this approach, the greedy algorithm, which is efficient and easy to implement, is chosen as the optimization method. The main principle of this method is that the contour nodes are moved one by one, and the total energy is minimized in a local area in each move. After several iterations, a local optimal solution can be obtained. In order to reduce the number of iterations and define a reasonable search range, the image resolution is adopted as the search step to move the nodes, and the search direction is determined with certain restrictions. As shown in Figure 5, the local optimal position of a contour node can be determined using the following steps:

1) One of the two segments that are connected to the node is chosen as the reference segment.

2) The local optimal position on this reference segment can be determined by first searching along the reference segment, and then searching along the direction perpendicular to the reference segment. The position that minimizes the total energy is selected as the updated position for the node.

3) Step 2) is repeated by choosing the other segment as the reference segment, and the plane coordinates of the local optimal position can be determined.

4) By searching along the vertical direction at the determined plane position, the optimal elevation for the node can be selected with energy minimization constraint.
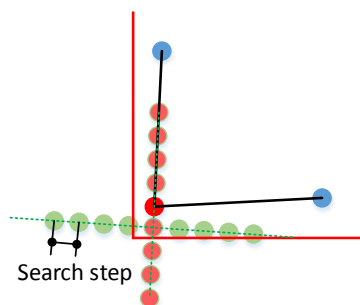

Figure 5. The search method for optimal position based on the reference segments. The black lines represent the initial contour, the red lines represent the actual contour. The red circle is the current position of the node, the green dashed lines are the search directions. The light red circles are the candidate positions when searching along the reference segment, the light green circles are the candidate positions when searching along the direction perpendicular to the reference segment.

3. EXPERIMENT

Three datasets located in vaihingen, Germany, which are provided by III/4 working group of ISPRS (Rottensteiner 2013), are used in the experiments to evaluate the performance of the proposed approach. Each dataset contains LiDAR point clouds and aerial images that are accurately registered. Figure 6 shows the experimental areas and the corresponding ground truths that are extracted by manual operation.
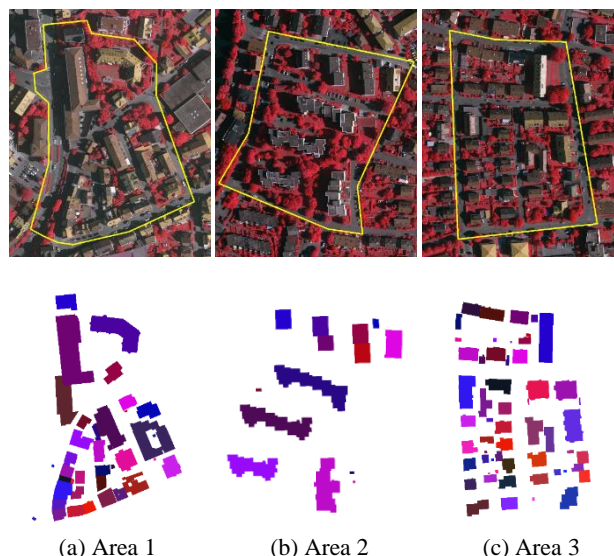

(a) Area 1   (b) Area 2   (c) Area 3

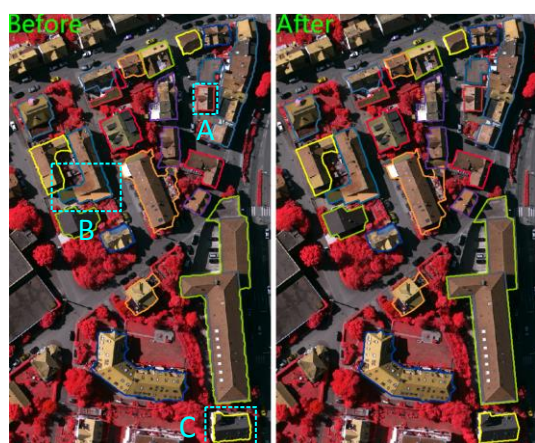Figure 6. The experimental datasets and their corresponding ground truths

The average density of LiDAR point cloud is 6.7 points / m^2, the resolution of aerial images is 8 cm. A classification method based on support vector machine (Zhang et al., 2013) is adopted to automatically detect the building points from LiDAR point cloud, and the errors in the classification results are modified by manual editing using TerraScan software. The initial building contours are extracted using alpha-shapes algorithm (Edelsbrunner and Mücke, 1994). Figure 7 shows the overlying views of Area 1 between the aerial images and the building contours before and after refinement, in which (a) shows the overall results, and (b) shows the enlarged views of the rectangular areas in (a). Figure 8-9 show the refinement results of Area 2 and 3, respectively. It can be seen that the building contours are obviously simplified and regularized, and fit better with the edge of buildings after refinement. In addition, it should be noted that some of contours are not strictly composed of rectangles, the proposed method achieves good results for building contours with certain complexity and non-rectangular corners.

In order to quantify the accuracy of the building contours before and after the refinement, by counting the number of true positive, false positive and false negative pixels with the ground truths, three kinds of indexes defined in (Lee et al., 2003), namely, completeness, correctness and quality are evaluated, respectively. Table 1 lists the statistics of these indexes for the three datasets. In overall, the quality indexes of the three datasets are both improved, which proves the effectiveness of the proposed method. The statistics also indicate that while the completeness indexes increase, the correctness indexes both decrease. This is largely because most of the refined building contours move outward from their initial positions.
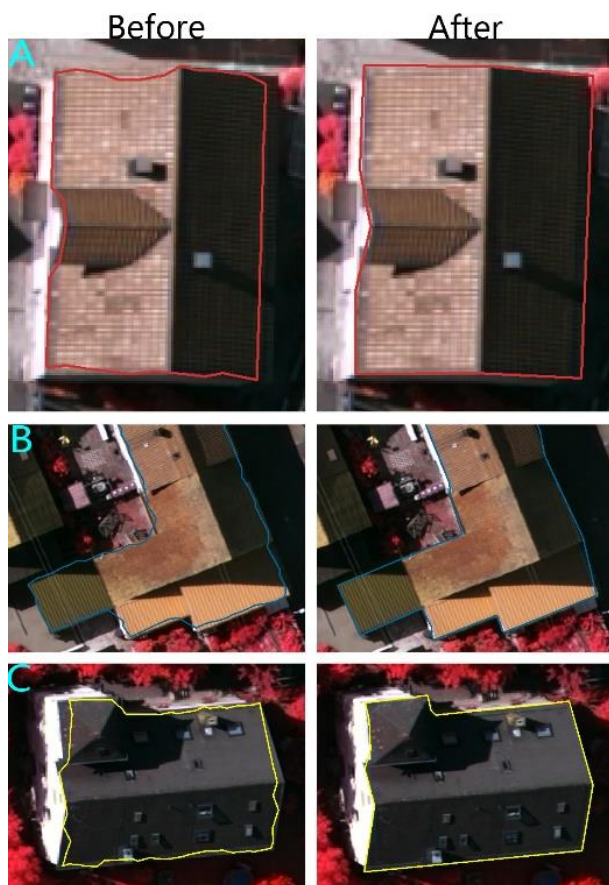
ID	Correctness (%)		Completeness (%)		Quality (%)	
	Before	After	Before	After	Before	After
Area 1	99.15	98.37	92.08	94.70	91.36	93.24
Area 2	99.46	98.63	93.07	95.10	92.61	93.86
Area 3	99.31	97.39	91.58	95.28	91.00	92.91
Average	99.31	98.13	92.24	95.03	91.66	93.34

Table 1. The evaluation results before and after the refinement

To find out whether the quality of a single building contour is improved or not, the evaluation for each individual building is performed in our approach. The statistics show that the quality indexes for most of the building contours are improved, the percentages of the building contours that have improved quality in three datasets are 84.0％，70.0% and 74.4％, respectively. In total, 77.0% of the building contours are improved after refinement. The processing times are 16, 10 and 17 s for three datasets, respectively. It can be considered that the efficiency of the proposed method has great application potential in the study of building contour extraction.
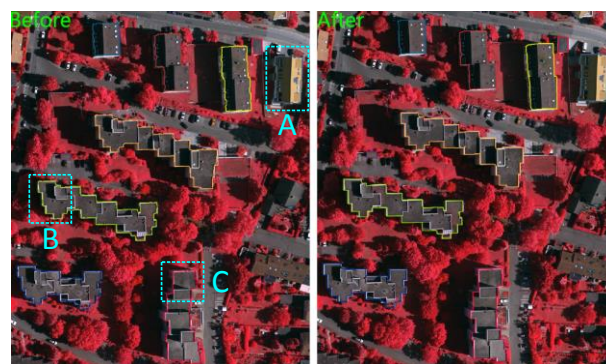

(a)

(b)

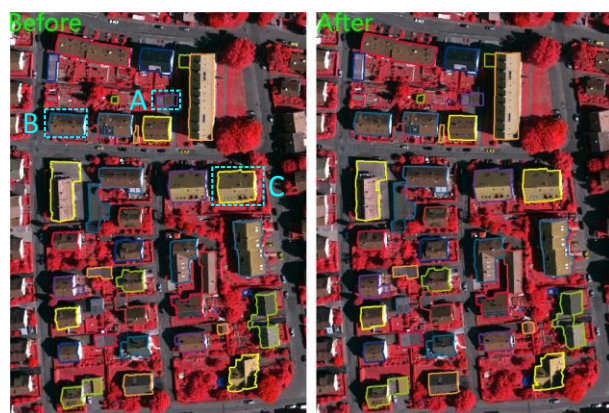Figure 7. The results of building contour refinement of Area 1
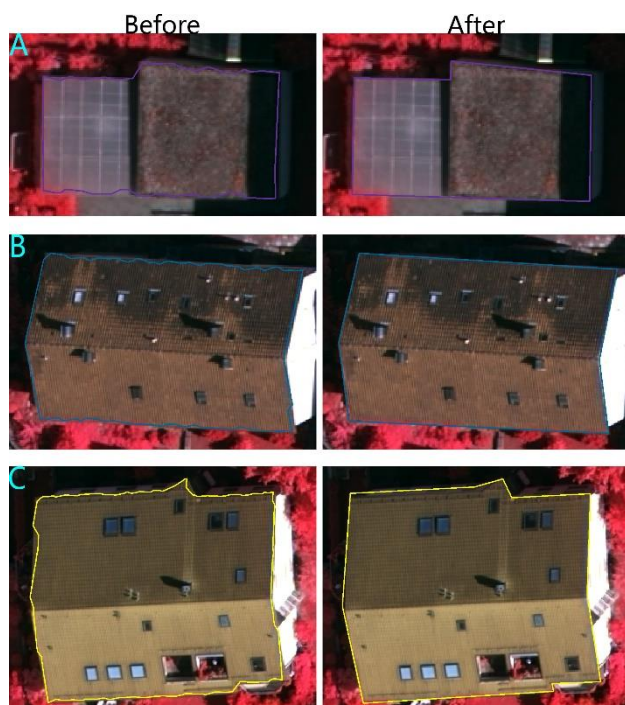

(a)

(b)

Figure 8. The results of building contour refinement of Area 2

(a)

(b)

Figure 9. The results of building contour refinement of Area 3

4. CONCLUSION

An improved Snake model is proposed to refine the LiDAR-derived building contours with the use of multi-view aerial images. In order to better meet the requirements of building contour refinement, the internal and external energy of Snake model are redefined, and the search method and the convergence criterion for the improved model are determined. Experiments on three datasets demonstrate that the quality indexes are both improved after refinement. The average quality index of the three datasets is improved from 91.66% to 93.34%. The evaluation results for every single building indicate that 77.0% of the total number of the building contours are updated with higher quality indexes. However, it should be pointed out that the interference from vegetation, shadow and the edge features inside the building contour may still result in failure of the proposed method. A subsequent study will investigate these problems in the future.

REFERENCES

Awrangjeb M., and Fraser C. S., 2014. Automatic segmentation of raw LiDAR data for extraction of building roofs. *Remote Sensing*, 6(5), pp. 3716-3751.

Awrangjeb M., Zhang C., and Fraser C. S., 2013. Automatic extraction of building roofs using LiDAR data and multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 83, pp. 1-18.

Bresenham J., 1977. A linear algorithm for incremental digital display of circular arcs. *Communications of the ACM*, 20(2), pp. 100-106.

Cheng L., Gong J., Li M., and Liu Y., 2011. 3D Building Model Reconstruction from Multi-view Aerial Imagery and Lidar Data. *Photogrammetric Engineering and Remote Sensing*, 77(2), pp. 125-139.

Dal Poz A. P., 2013. Photogrammetric refinement of laser-based building roof contours. In: *Proceedings of Urban Remote Sensing Event, 2013 Joint*, pp. 29-32.

Dorninger P., and Pfeifer N., 2008. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. *Sensors*, 8(11), pp. 7323-7343.

Douglas D. H., and Peucker T. K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10(2), pp. 112-122.

Edelsbrunner H., and Mücke E. P., 1994. Three-dimensional alpha shapes. *ACM Transactions on Graphics*, 13(1), pp. 43-72.

Fazan A. J., Dal Poz A. P., 2013. Rectilinear building roof contour extraction based on snakes and dynamic programming. *International Journal of Applied Earth Observation and Geoinformation*, 25, pp. 1-10.

Galvanin E. A. D. S., and Dal Poz A. P., 2012. Extraction of building roof contours from LIDAR data using a Markov-Random-Field-based approach. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3), pp. 981-987.

Lee D. S., Shan J., and Bethel J. S. 2003. Class-guided building extraction from IKONOS imagery. *Photogrammetric Engineering & Remote Sensing*, 69(2), pp. 143-150.

Li G., An J., Wan Y., 2012. Automatic building extraction based on improved watershed segmentation, mutual information match and snake model. *International Journal of Computer Applications in Technology*, 43(2), pp. 147-154.

Li H., Zhong C., Hu X., Xiao L., and Huang X., 2013. New methodologies for precise building boundary extraction from LiDAR data and high resolution image. *Sensor Review*, 33(2), pp. 157-165.

Mongus D., Lukač N., and Žalik B., 2014. Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, pp. 145-156.

Niemeyer J., Rottensteiner F., and Soergel U., 2014. Contextual classification of LiDAR data and building object detection in urban areas. *ISPRS journal of photogrammetry and remote sensing*, 87, pp. 152-165.

Peng J., Zhang D., and Liu Y., 2005. An improved snake model for building detection from urban aerial images. *Pattern Recognition Letters*, 26(5), pp. 587-595.

Rottensteiner F., Sohn G., Gerke M., and Wegner J. D., 2013. ISPRS test project on urban classification and 3D building reconstruction. *Commission III-Photogrammetric Computer Vision and Image Analysis, Working Group III/4-3D Scene Analysis*, pp. 1-17.

Zhang J., Lin X., and Ning X., 2013. SVM-based classification of segmented airborne LiDAR point clouds in urban areas. *Remote Sensing*, 5(8), pp. 3749-3775.

Zhao Z., Duan Y., Zhang Y., and Cao R. 2016. Extracting buildings from and regularizing boundaries in airborne LiDAR data using connected operators. *International Journal of Remote Sensing*, 37(4), pp. 889-912.