EXTRACTION OF PLANES CONSISTING OF BUILDINGS FOR MODELING USING CLOSE-RANGE PHOTOGRAMMETRY

Yuta Kurokawa*[1] and Junichi Susaki[2]

[1]Graduate student, Department of Civil and Earth Resources Engineering, Kyoto University,
C1-1-209, Kyotodaigaku-katsura, Nishikyoku, Kyotoshi, 615-8540, Japan; TEL:81-75-383-3302;
E-mail: y.kurokawa@ky3.ecs.kyoto-u.ac.jp

[2]Professor, Department of Civil and Earth Resources Engineering, Kyoto University,
C1-1-206, Kyotodaigaku-katsura, Nishikyoku, Kyotoshi, 615-8540, Japan; TEL:81-75-383-3300;
E-mail: susaki.junichi.3r@kyoto-u.ac.jp

KEY WORDS: close-range digital photogrammetry, dense urban areas, passpoints, modeling, region growing

ABSTRACT: In recent years, the demand of three-dimensional model of buildings in cultural heritage and urban areas is growing rapidly. In dense urban areas, close-range digital photogrammetry is appropriate in terms of the cost and portability. However, in dense urban areas, the orientation may fail especially in case with narrow streets and many obstacles. In addition, it is quite time-consuming to manually select passpoints necessary for building modeling. Therefore, the objective of this study is to examine the preprocessing for the efficient modeling of the buildings in dense urban areas using digital photogrammetry. Automatic extraction of planes consisting of a building can help users in modeling. In the proposed approach, the planes are extracted using the three-dimensional coordinate and color data of points on buildings. First, relative orientation is conducted and a large number of passpoints are generated using SIFT (Scale-Invariant Feature Transform). Then, the passpoints are automatically classified into several independent planes. Finally, the boundary of the plane is determined through region-growing algorithm by referring to the brightness of images. The plane classification prior to the region-growing can merge the shadowed regions with non-shadowed regions, and reduce the shadow effect. In the experiment, images were obtained in Hanoi Old Quarter. As a result, it was found that three or four planes, walls and commercial signs, were successfully extracted from these images by using the proposed method. On the other hand, the other planes were not extracted because the planes had been divided by the obstacles or the patterns on the planes were intricate. In future, it is necessary to consider how to improve the accuracy of the extraction using the three-dimensional coordinate data instead of the color data.

1. INTRODUCTION

In recent years, the demand of three-dimensional model of buildings in cultural heritage and urban areas is growing rapidly. For example, in Southeast Asia, many areas come to be rapid urbanization. Hanoi, Vietnam is one of the fastest growing cities, and is also historic city because of its having about 2,000 historic sites and buildings. In this area, the three-dimensional model of the urban area is helpful in the landscape conservation and the disaster management. In generally, the proper measuring method of modeling is selected according to the purpose and the target area. In dense urban areas, such as Hanoi, close-range digital photogrammetry is appropriate in terms of the cost and portability. With the development of bundle adjustment method, close-range digital photogrammetry is developed; it was possible to measure with high accuracy by using a commercial digital camera. However, in the images taken at the Hanoi Old Quarter, there are many obstacles such as trees and bikes on the street. They often cover the plane of the target building. Therefore it is difficult to measure in this situation. Furthermore, when modeling by using close-range digital photogrammetry, the corresponding parts between two images (typically, the end points of the building surface) have to be extracted to obtain the information of building such as width and height. This work is often done manually. Therefore, if an obstacle is in the images, it takes a lot of effort. In essence, it is very useful to extract the plane containing the building automatically in spite of obstacles. In addition, the images taken at the Hanoi Old Quarter, there are many signs and windows on the front and the side wall of the building. Each of them has a plane parameter which is different from the other parameters. The plane object like this makes a small area in the image. When taking by using two cameras, each corresponding areas between two images are associated with a different projection transformation matrix. Therefore, the extraction of multiple regions of small areas in the images leads to determine some projection transformation matrixes accurately (Sakamoto, 2003) (Ueshiba, 2002) (Xu, 2000) (Zhang, 1995). This is an important technology to improve the accuracy of its three-dimensional reconstruction. Based on the above, the purpose of this study is to accurately extract multiple regions of small areas on the building in the images. In the proposed method, the contours around the regions are extracted by using two constraints; the three-dimensional geometric conditions and the color information of the images.

2. THE PROPOSED METHOD

In this method, first, the feature points are extracted by using SIFT (Scale-Invariant Feature Transform) on each image separately. SIFT is one of the operators to extract the feature points (Lowe, 2004). And then, the association program makes the extracted feature points corresponded with each other (Fujiyoshi, 2007). After the association of corresponding parts, the three-dimensional feature points coordinates is obtained. Second, each point is classified according to the plane on which it exists in the terms of the geometry that points make up planes. Next, in one image, the classified points are defined as the starting points, and the region of the plane starts to be extended by using the color information of the image plane. Finally, the area is determined by extracting the boundaries.

2.1 Extraction of passpoints by using SIFT

In this study, on the images, a large number of corresponding points (passpoints) are obtained in term of the technology of automatic extraction of corresponding points by using the SIFT. After the relative orientation by using this passpoints, their three-dimensional coordinates are obtained by stereo matching.

2.2 Classification of three-dimensional point cloud

The automatic classification of three-dimensional point cloud is divided into the following three tasks significantly. Through these procedures, each point is classified according to the plane on which it exists and is allocated the plane number.

(1) Extraction of sub-planes: First, it is examined how many points there are within 2 m radius around any point in three-dimensional point cloud data. If there are more than 5 points, the candidate plane (sub-plane) is made up by using the least squares method sequentially as shown in Figure 1.

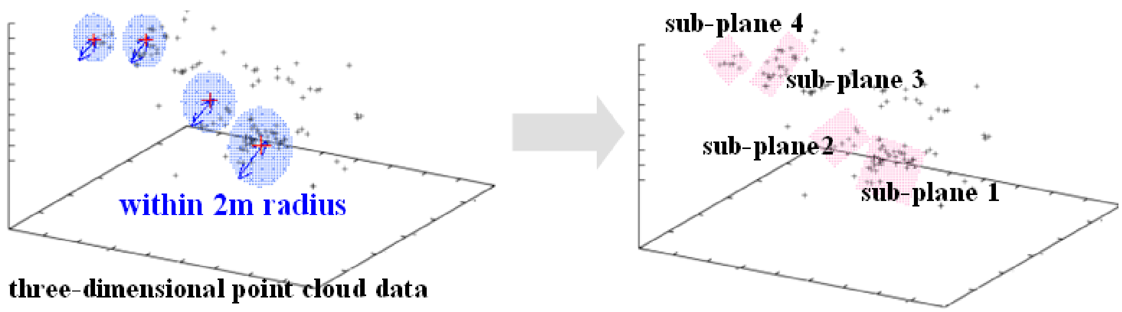

Figure 1 Extraction of sub-plane

The equation of sub-plane [i] (i = 1,2, ...) is defined by using the dimensional coordinates (x, y, z) as follow respectively.

$$a_i x + b_i y + c_i z + d_i = 0 \quad (i = 1, 2, \cdots) \tag{1}$$

(2) Integration of sub-planes: Next, several sub-planes are examined whether they share the same plane. If they share the same plane, they are integrated as one plane. In this study, two conditions are defined. One is the condition about the distance from one centroid of the points that make up the sub-plane to the other centroid. In this condition, the maximum distance was set 2 m. The other condition depends on whether the centroid of one sub-plane (for example, sub-plane 1) satisfies the equation of the other sub-plane (for example, sub-plane 2) within the error of 0.2 m. If these two conditions are satisfied, the integrated plane is configured by using the all points that make up two sub-planes according to the procedure of (1). Then, the independent planes are finally generated as shown in Figure 2.

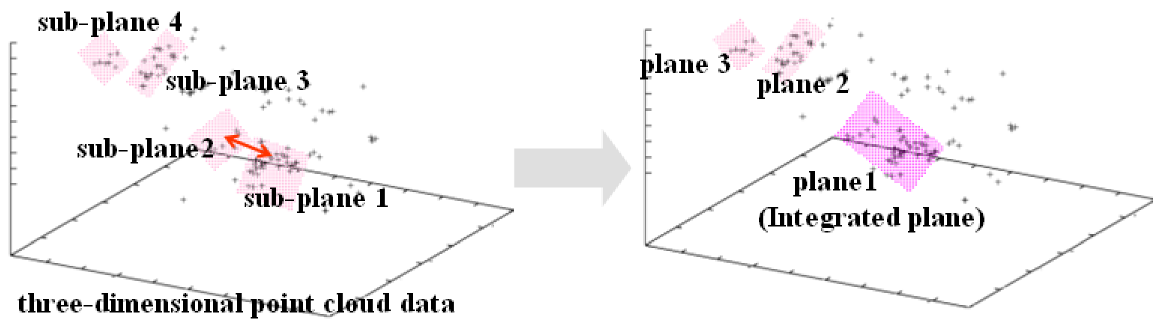

Figure 2 Integration of sub-plane

(3) Integration of unclassified points: Finally, the points unclassified into any sub-plane in the procedure of (1) are examined whether they are classified. The result of this final classification depends on the distance from one unclassified point to a sub-plane. If the distance is less than 0.2 m, the point is classified.

2.3 Region-growing by using color information

In this section, the method of region-growing is introduced. Region-growing is proposed by Haralick and Shapiro in 1985 (Shapiro, 2001) and one of the methods that expand the region of an area on images under some condition. Region-growing is a simple region-based image segmentation method. It is also classified as a pixel-based image segmentation method since it involves the selection of starting points. This approach to segmentation examines neighboring pixels of starting points and determines whether the pixel neighbors should be added to the region. The process is iterated on, in the same manner as general data clustering algorithms. The algorithm in this study is consisted of two procedures described below. The color information is used as the condition of extending an area on images.

(1) For pixels in eight directions around a starting point, their RGB values is compared with that of the starting point respectively. If the deference is within $\pm$ 4, the neighboring pixels are determined to belong to the same plane which the starting point belongs to.

(2) The new starting points are selected from the pixels determined to belong to the same plane. The work of (1) is repeated until the deference is not within $\pm$ 4.

In this study, initial starting points are the three-dimensional points data classified according to the method of chapter 2.

2.4 Extraction of boundary

After the planar region is confirmed by using the region-growing method described in section 2.3, the boundary of the planar region often contains irregularities. Therefore, it is demanded to extract the boundary of the plane as a straight line. In this section, the method to approximate a border to a straight line is introduced. This method is based on Hough transform.

(1) Hough transform: The Hough transform is a feature extraction technique used in image analysis, computer vision, and digital image processing (Shapiro, 2001). The purpose of the technique is to find imperfect instances of objects within a certain class of shapes by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform. The theory is described bellow. Any point (x, y) is in the two-dimensional coordinate system. All of the lines passing through the point (x, y) are determined by two parameters (ρ, θ). ρ is the distance from the origin perpendicular to the straight line. And θ is angle between the x-axis and the segment of the distance ρ. It is shown as the following equation.

$$\rho = x\cos\theta + y\sin\theta \tag{2}$$

(2) Extraction of lines by using Hough transform: After the region-growing processing of section 2.3, the planar region is obtained as shown in Figure 3. In this stage, the boundaries are uneven. Therefore, they are demanded to be approximated to straight lines. First, Hough transform is applied to the points which form the boundary in order to find the intersection (ρ_1, θ_1) which most curves pass through in ρ-θ coordinate space. And then, the second most intersection (ρ_2, θ_2) is determined. θ has to stratify the following condition since the line of the boundary is horizontal or vertical for x-y coordinate system of an image.

$$\begin{gathered} -5^\circ \le \theta \le 5^\circ \quad or \\ -90^\circ \le \theta \le -85^\circ, 85^\circ \le \theta \le 90^\circ \end{gathered} \tag{3}$$

Next, the lines of l_1, l_2 are made by substituting the parameters of (ρ_1, θ_1), (ρ_2, θ_2) for equation (2) respectively. These lines' length is limited by the planar region. l_1 and l_2 are shown in Figure 4. Finally, two missing lines are interpolated and the final boundary is formed as shown in Figure 5.

Figure 3 Initial boundary estimated by using region growing

Figure 4 Two lines extracted
by using Hough transform

Figure 5 Final boundary extracted
by the proposed method

3. EXPERIMENT

This method of extracting planar region was examined for the images taken in Hanoi Old Quarter bellow.

3.1 Experiment by using the different images in having pass points

In this experiment, two data sets were used. One of the sets included the two images having 50 passpoints extracted by using SIFT. The other included the two images having 195 passpoints. These are shown in Figure 6 and 7 respectively. And the accuracy of these orientations is shown in Table 1. These orientations were carried out by using the software "Image Master" made by Topcon. When the proposed method was applied to these images, the result of classification of the three-dimensional points data were shown in Figure 8 and 9, and the extracted regions were shown in Figure 10and 11 respectively. As this result, it was revealed that the planar regions were extracted a little more accurately in the images concluded a lot of pass points than a few pass points. Specifically, the plane in left side of images was extracted more wildly in Figure 7 than in Figure 6.

Figure 6 Data set (50 pass points)

Figure 7 Data set (195 pass points)

Table 1 Accuracy of orientation

Pass Points	Y-parallax	RMSE
50	0.30	0.15
195	0.21	0.11

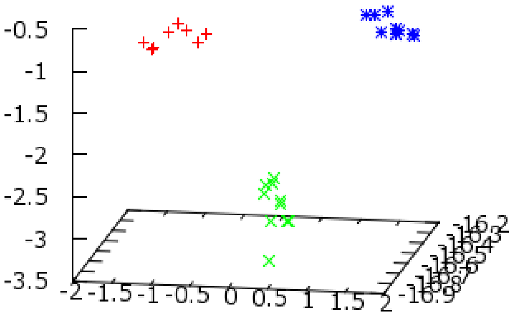

Figure 8 Classification of the 50 points data

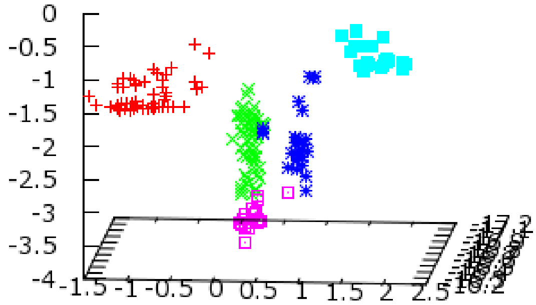

Figure 9 Classification of the 195 points data

Figure 10 Planar region obtained by using the proposed method (50 pass points)

Figure 11 Planar region obtained by using the proposed method (195 pass points)

3.2 Experiment by using the images taken diagonally

In this experiment, the proposed method was applied to the images taken diagonally as the face of building in an image was not parallel to the same face in the other image. The used data set included three successive images as shown in Figure 12. And there were 50 passpoints per pair of neighboring images. The accuracy of the orientation was shown in Table 2. When the proposed method was applied to these images, the result of classification of the three-dimensional points data were shown in Figure 13, and the extracted regions were shown in Figure 14. It was possible to extracting the boundary. In generally, the orientation of these images is difficult since they are able to overlap well. Therefore, it is considered that the three-dimensional reconstruction is accurately obtained by using some projection transformation matrixes of multiple regions extracted in this method. However, the extraction planar regions is difficult by the complex pattern on the plane and the difference in the brightness between the forward and the backward plane of the images. Specifically, the plane in right side of images was not extracted perfectly as shown in Figure 14.

Figure12 Data set

Table 2 Accuracy of orientation

Pair	Y-parallax	RMSE
Pair1	0.40	0.68
Pair2	0.23	0.27

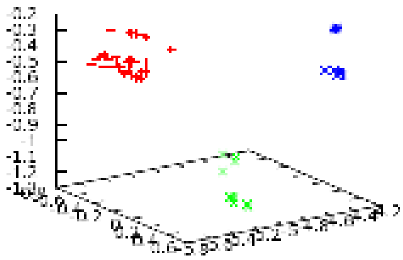

Figure 13 Classification of the points data

Figure 14 Result of region growing (3 images)

4. CONCLUSION

In this study, the combination of three and two dimensional data makes it possible to extract the corresponding planar regions accurately on the images. It is difficult to configure the planar region strictly by using only the three-dimensional points data (Kawakami, 2005). On the other hand, by using only the two-dimensional image data, the planar region is often separated due to the obstacle on the image and extracted unnecessary area like shadowed regions. Therefore, this method is effective to extract the planar regions strictly. In this study, it is the premise that passpoints are extracted accurately by using SIFT. However, passpoints extracted by the method using SIFT are not necessarily correct and they are often concentrated into a part of an image. Therefore, it is demanded to establish the method of extracting passpoints accurately and sparsely on the images. If there are a lot of accurate passpoints evenly throughout the image, it is possible to extract the planar region accurately. Furthermore, the effect of obstacles is not removed perfectly since the region growing is carried out in two-dimensional space. Therefore, the obstacles are able to be removed as the subjects on the foreground by using the region growing in the three-dimensional space. These proposed methods have to be examined in the future study.

REFERENCES

Kawakami, Y., Y. Ito and Y. Kanazawa, 2005, A robust method for detecting planar regions based on random sampling using distributions of feature points. IEICE Transactions, D-Ⅱ, Voll. J88-D-Ⅱ, No.2, pp. 313-324

Sakamoto, T., Y. Kawakami and Y. Kanazawa, Jan, 2003. Robust 3D reconstruction by plane region detection. Gakken Information Report, CVIM-136-24, pp.179-186.

Ueshiba, T., F. Tomita, Nov, 2002. Multiple camera calibration system using planar patterns. Gakken Information Report, CVIM-135-9, pp.47-54.

Xu, G., J. Terai and H.-Y. Shum, June, 2000. A linear algorithm for camera self-calibration, motion and structure recovery for multi-planar scenes from two perspective images. Proc, Comput, Vision Patt, Recog, pp.Ⅱ:474-479. South Carolina, U.S.A.

Zhang, Z., R. Deriche, O. Faugeras, and Q.-T. Loung, Oct, 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artif, Intell., vol.78, pp.87-119.

Lowe, David G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60, 2, pp. 91-110.

Fujiyoshi, H., 2007. Extraction of feature based on Gradient-HOG and SIFT-. Information Processing Society, Research Reports, CVIM 160, pp. 211-224.

Shapiro, Linda G., George C. Stockman, 2001. Computer Vision. New Jersey, Prentice-Hall.

ACKNOWLEDGEMENT

This research was supported by a grant from Japan Construction Information Center.