AUTOMATIC SELECTION OF PASSPOINTS IN DIGITAL CLOSE RANGE PHOTOGRAMMETRY IN DENSE URBAN AREAS

Yuki Ohta[*a] and Junichi Susaki[b]

[a] Student, Department of Civil and Earth Resources Engineering, School of Engineering, Kyoto University, C1-1-209, Kyotodaigakukatsura , Nishikyo-ku, Kyoto 615-8540, Japan; Tel/Fax: 075-383-3302
E-mail: yuki-ota@nishinomiya.mbox.media.kyoto-u.ac.jp

[b] Associate Professor, Department of Civil and Earth Resources Engineering, School of Engineering, Kyoto University, C1-1-206, Kyotodaigakukatsura , Nishikyo-ku, Kyoto 615-8540, Japan; Tel/Fax: 075-383-3300
E-mail: susaki.junichi.3r@kyoto-u.ac.jp

KEY WORDS: Digital close-range photogrammetry, Three dimensional city modeling, Dense urban areas, Automatic selection of passpoints

ABSTRACT: These years, the demand for three-dimensional (3D) urban building model is growing. Especially in dense urban areas such as Preservation District in Kyoto, the efficient and low-cost 3D modeling for landscape analysis and predicting fire spread is required. Digital close-range photogrammetry has potential to meet the demand for efficient and low-cost 3D modeling in dense urban areas and the method for automatic selection of passpoints is required to develop the efficiency. In present research, the authors developed the method for automatic selection of passpoints. In this method, matches on stereo-pairs are extracted using Scale-Invariant Feature transform (SIFT) (Lowe, 2004), and matches on back/foreground and erroneous matches are removed using the image-coordinates of matches because that points cause to decrease the accuracy of orientation. Furthermore, the method for reducing the number of passpoints is developed to reduce the computation time of orientation. Proposed method was adapted to 17 successive stereo-pairs taken in dense urban areas in Kyoto, and it was revealed that matches on fore/background and erroneous matches could be removed and the number of passpoints could be reduced, preventing matches from remaining in small area on image. As a result, the accuracy of orientation was acceptable and the computation time was short. In near future, absolute orientation will be performed using the measurement data taken with Total Station and the accuracy of the measurement with photogrammetry using the passpoints selected with this method will be verified in detail.

1. INTRODUCTION

These years, the demand for three-dimensional (3D) urban building model is growing. Especially in dense urban areas such as Preservation District in Kyoto, the efficient and low-cost 3D modeling for landscape analysis and predicting fire spread is required. Using digital close-range photogrammetry, detailed data of side surface of building can be measured. Furthermore, the field survey can be taken efficiently, and the digital camera is not expensive. Therefore, digital close-range photogrammetry has potential to meet the demand for efficient and low-cost 3D modeling in dense urban areas. However, it was problem that manual selection of passpoints requires a lot of time. When the measurement is taken in dense urban areas, many successive stereo-pairs cause to inefficiency in manual selection of passpoints. The method for automatic selection of passpoints is required to develop the efficiency in modeling in dense urban areas.

There are various image-matching methods which have potential to be adapted to automatic selection of passpoints. However, the results of image-matching always include erroneous matches and the algorithm to remove erroneous matches is required. In photogrammetry software, erroneous matches are removed if the points have large residual error s after orientation. However, using this conventional function, it takes a lot of time to calculate to select the correct passpoints because the orientation has to be performed again and again. Furthermore, there are some matches on foreground such as trees and fences, and on background on images taken in dense urban areas. When the models of the objects on fore/background are not generated, the matches on fore/background also have to be removed though they are correct matches.

In present research, the authors developed the method for automatic selection of passpoints of successive stereo-pairs. In this method, matches on stereo-pairs are extracted using Scale-Invariant Feature transform (SIFT) (Lowe, 2004), and matches on back/foreground and erroneous matches are removed using the image-coordinates of matches. In section 3, the algorithm of the proposed method is explained and the result is shown when the method was adapted to 18 successive images of buildings in dense urban area in Kyoto.

2. STUDY AREA

The area near Kodai-temple in Higashiyama-ku, Kyoto (Figure 1) was selected as study area. This area is in Sanneizaka Preservation District and it has plenty of traditional houses and landscape. 3D building models for landscape analysis and predicting fire spread is required. 18 successive images of 8 lined buildings were taken for modeling. The images were taken using digital camera (Eos Kiss X3, Canon).

Figure 1 Buildings near the Kodai-temple

3. PROPOSED METHOD FOR AUTOMATIC SLECTION OF PASSPOINTS

Proposed method for automatic selection of passpoints is composed of following 3 steps.
(1) Extraction of matches on stereo-pairs using Scale-Invariant Feature Transform (SIFT)
(2) Removing matches on fore/background and erroneous matches
(3) Reducing the number of matches

First, matches are extracted on each stereo-pair using SIFT. Then, extracted matches include erroneous matches and they have to be removed. Matches on fore/background also have to be removed because only buildings facing the road are measured in this method for efficient modeling. If fore/background models are generated, more images are required and the calculation time of orientation increase. Figure 2 shows example of matches on fore/background and erroneous matches. Finally, the number of matches is reduced in order to decrease the calculation time of orientation for many stereo-pairs. Step (2) and (3) are explained in the following subsections.

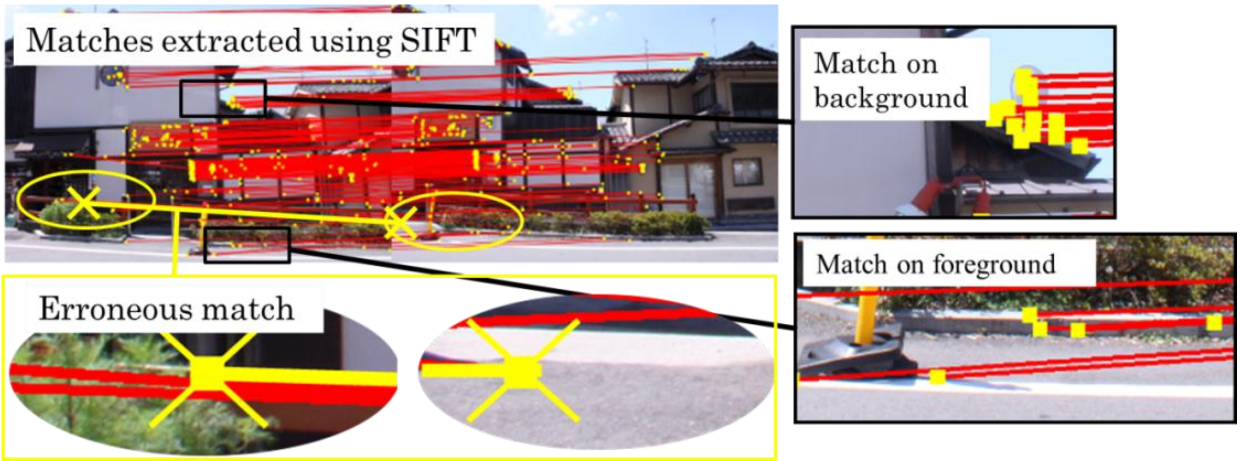

Figure 2 Matches extracted using SIFT, matches on fore/background and erroneous matches

3.1 Removing matches on fore/background and erroneous matches

Extracted match Pi has coordinates (x_{left-i}, y_{left-i}) on left image and $(x_{right-i}, y_{right-i})$ on right image. The values of dx_i, dy_i and d_i are calculated as follows and these values are used in following procedure.

$$dx_i = x_{left-i} - x_{right-i}$$

$$dy_i = y_{left-i} - y_{right-i}$$
$$d_i = \sqrt{dx_i^{\,2} + dy_i^{\,2}}$$

3.1.1 Removing matches on fore/background: The nearer point P_i is from camera, the larger the values of d_i become (Figure 3). Therefore, point P_i is likely to be on foreground if the value of d_i is large, and to be on background if the value of d_i is small. Matches on fore/background can be removed, considering the relation between the value of d_i and the distance to camera. First, a frequency distribution of the value of d_i is represented. Then, area where the most number of the value of d_i is counted is detected. The area has a certain width and the width is set to 300 through the empirical examination. After that, the average (μ_d) and the standard deviation (σ_d) of the value of d_i in the area are calculated. Finally, P_i is removed as matches on fore/background if following inequalities are satisfied. Figure 4 shows the result that this procedure was adapted to matches on stereo-pair img3-img4.

$$d_i < \mu_d - \sigma_d \quad (P_i \text{ is removed as matches on background})$$
$$d_i > \mu_d + \sigma_d \quad (P_i \text{ is removed as matches on foreground})$$

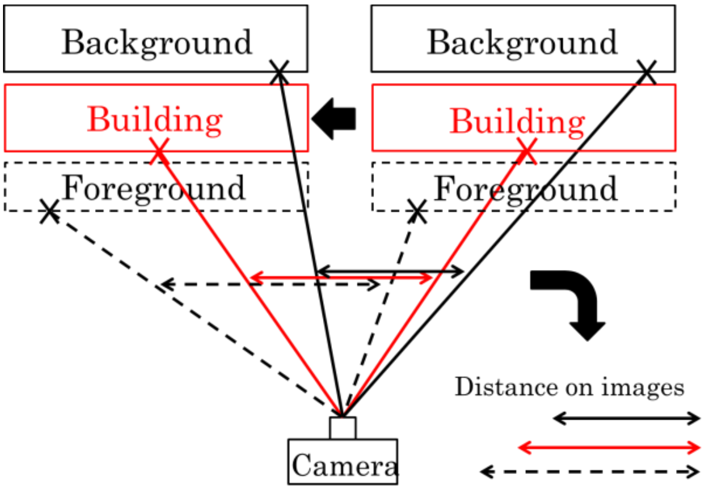

Figure 3 Relation between the value of d_i and the distance to camera.

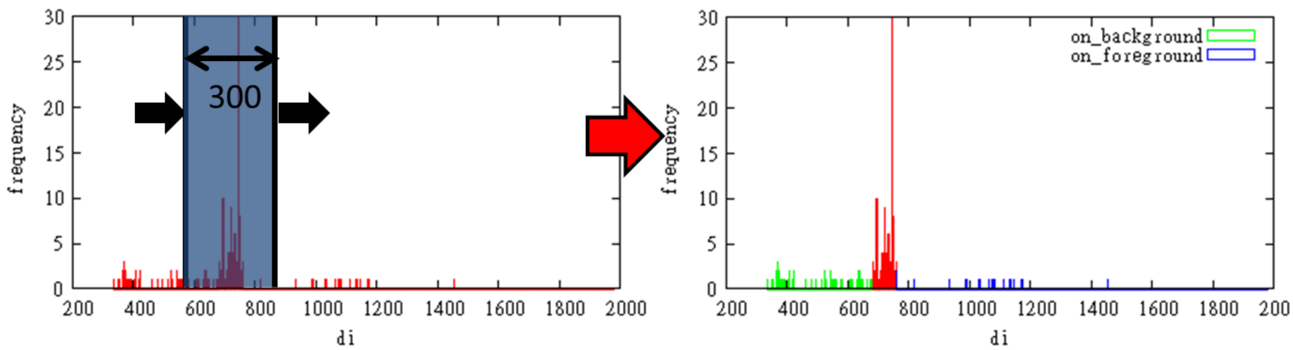

Figure 4 Removing matches on fore/background (img3-img4)

3.1.2 **Removing erroneous matches**: Because matches on fore/background are removed in 3.1.1, it is assumed that all remaining matches are on side walls of buildings. Assuming that all remaining matches are on the same plane, correct matches can be removed in the following procedure using the value of dx_i and dy_i of matches.

First, the values of x_{left-i}, y_{left-i} and dx_i of the remaining matches are plotted. Then, regression plane of the point cloud is estimated using least-squares method. When the values of x_{left-i}, y_{left-i} and dx_i of correct matches are plotted, it may be assumed that the point cloud is on or near the regression plane. Next, each distance between match and the regression plane are calculated ($dist\_dx_i$), and the average (μ_{dist_dx}) and the standard deviation (σ_{dist_dx}) of the distance are calculated. After that, $dist\_dy_i$, μ_{dist_dy} and σ_{dist_dy} are calculated using the values of x_{left-i}, y_{left-i} and dy_i in same procedure. Finally, if following inequalities are satisfied, it is judged that the point is too far from the regression plane and it is removed as erroneous matches.

$$dist\_dx_i > \mu_{dist\_dx} + \sigma_{dist\_dx}$$
or
$$dist\_dy_i > \mu_{dist\_dy} + \sigma_{dist\_dy}$$

Figure 5 was the result that this procedure was adapted to matches on stereo-pair img3-img4.

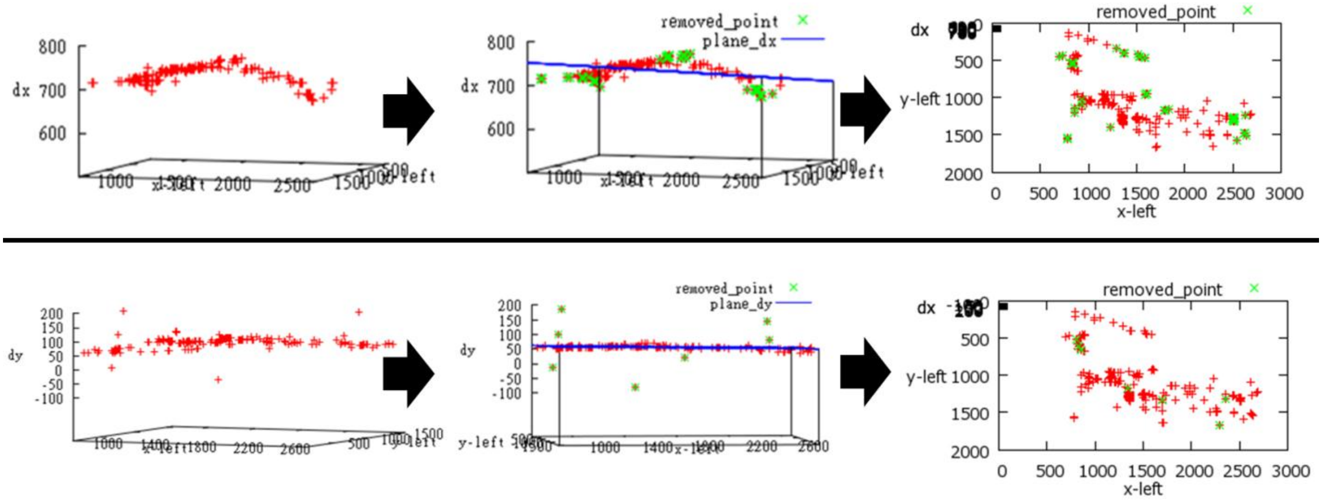

Figure 5 Removing erroneous matches (img3-img4)

3.2 Reducing the number of matches

Then, a lot of matches remain and it takes a lot of time to perform the orientation if all remaining matches are used as passpoints. Therefore, the number of matches has to be reduced to decrease the computation time of orientation. Preventing matches from remaining in small area on image, the number of matches is reduced in following process.

(1) A match is selected as passpoint (Passpoint-0) if its distance of SIFT descriptors is smallest.
(2) The match is removed if the distance to Passpoint-0 is less than a threshold *r* (pixels). The match is selected from remaining matches if its distance of SIFT descriptors is smallest (Passpoint-1). Passpoint-2, Passpoint-3…are selected in same process.
(3) The process is stopped if there is no match to be selected. Then, if the number of passpoints is less than a threshold, this process is done again from (2), reducing the value of *r*.

Initial value of *r* was set to 500 through the empirical examination. SIFT descriptor is 128 dimensional vector, and the shorter the distance of two points constituting match is, the more similar their features are (Lowe, 2004). Therefore, the distance of SIFT descriptors was used to select more reliable matches.

Proposed method for automatic selection of passpoints was adapted to 17 successive stereo-pairs. Table 1 and Figure 6 show the results. The results show that matches on fore/background and erroneous matches were removed. Furthermore, the results also show that the number of matches was reduced and the remaining matches were not in small area on image.

Table 1 The number of keypoits and matches after each procedure

	img1	img2	img3	img4	img5	img6	img7	img8	img9	img10	img11	img12	img13	img14	img15	img16	img17	img18
Keypoints extracted with SIFT	12353	11333	14487	13030	12978	12626	14507	16018	14539	14646	14158	21976	21744	15019	13696	20369	22649	18807
Matches before removing	690	607	339	159	296	221	421	365	498	462	135	356	509	873	614	680	490	
Matches after removing	338	398	204	45	116	124	232	187	249	211	36	191	284	434	377	413	193	
Matches after reducing	10	11	11	10	11	12	10	10	10	10	10	15	10	14	10	10	10	

Figure 6 Passpoints of 17 successive stereo-pairs selected automatically using proposed method

Remaining matches are selected as passpoint and orientation is performed using these passpoints. The result of orientation for 17 successive stere-pairs is shown in next section.

4. RESULT OF ORIENTATION FOR SUCCESSIVE STEREO-PAIRS

Using passpoints selected in Section 3, orientation for 17 successive stereo-pairs was performed using Image Master (Topcon). The results of orientation and computation time are shown in Table 2 and 3, respectively. Size of images used in present research is 2800x1867 (pixel-pixel) and the both of *x* and y resolution are 0.0046 mm.

Table 2 Result of orentation for 17 successive stereo-pairs

	img1-img2	img2-img3	img3-img4	img4-img5	img5-img6	img6-img7	img7-img8	img8-img9	img9-img10	img10-img11	img11-img12	img12-img13	img13-img14	img14-img15	img15-img16	img16-img17	img17-img18
y-parallax (pixel)	2.48	2.91	2.28	1.9	2.52	3.06	2.37	2.68	1.7	1.79	1.4	1.43	0.74	2.03	3.14	0.77	1.68
residual error (pixel)	1.31	1.43	1.43	1.06	1.49	1.49	1.46	1.28	1.1	0.82	0.7	0.7	0.77	0.98	1.01	1.01	0.67

Table 3 Computation time of each procedure

Keypoint extraction using SIFT	Keypoint matching	Removing	Reducing	Orientation
Approx. 5min	Approx. 10min	Approx. 1.5sec	Approx. 2sec	Approx. 6sec

Table 2 shows that the y-parallaxes and residual errors of each passpoint were not large and orientation was successful. Table 3 shows that while it took approximately 15 minutes to extract keypoints and matches, it took only approximately 10 seconds to remove matches on fore/background and erroneous matches and to reduce the number of matches. In present research, expected accuracy tolerance of distance was approximately 30 cm, considering the accuracy of digital maps used in other researches in predicting fire spread (Minami et al), for example Digital Maps 2500 (Geospatial Information Authority of Japan). Then, the accuracy of orientation was acceptable.

5. CONCLUSIONS

In this paper, the method for automatic selection of passpoints of successive stereo-pairs in dense urban area was proposed. Using the value of image coordinates of matches, matches on fore/background and erroneous matches could be removed from matches extracted with SIFT, and the computation time was short. Furthermore, the number of passpoints was reduced, preventing matches from remaining in small area on image. As a result, the orientation for 17 successive stereo-pairs could be performed efficiently and the accuracy was acceptable. In near future, absolute orientation will be performed using the measurement data taken with Total Station and the accuracy of the measurement with photogrammetry using the passpoints selected with this method will be verified in detail.

ACKNOWLEDGHMENT

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

REFERRENCE

David G. Lowe, 2004. Distinctive Image Features from Scale-Invariant Keypoints. Interntional Journal of Computer Vision, 60, 2, pp. 91-110.

S. Minami, M. Tomatsu, S. Takeuchi, K. Hayashi, M. Michiwaki Y. Ohno, Y. Higuchi, Development of Earthquake Disaster Mitigation Maps for Municipalities, AIJ J Technol. Dee. No.9, 261-264, Dec, 1999.

Geospatial Information Authority of Japan, Digital aerial photogrammetry, Manual of public surveying, Technical material of Geospatial Information Authority of Japan, A-1-No.314, 2006.