AUTOMATIC TIE-POINT MATCHING USING SURF-BASED IMAGE MATCHING SCHEME

Jyun-Ping Jhan[1], Jiann-Yeou Rau[2], Chien-Chuan Lin[3] and Yi-Chen Shao[4]
[1]Graduate Student, Department of Geomatics, National Chung Kung University,
No.1, University Rd., Tainan City 701, Taiwan
E-mail: riddle0104@hotmail.com

[2]Assistant Professor, Department of Geomatics, National Chung Kung University,
No.1, University Rd., Tainan City 701, Taiwan
E-mail: jyrau@mail.ncku.edu.tw

[3] Ph.D Candidate, Department of Engineering Science, National Chung Kung University,
No.1, University Rd., Tainan City 701, Taiwan
roger.lincc@msa.hinet.net

[4] Associate Professor, Department of Civil Engineering, China University of Science and Technology
No.245, Sec. 3, Academia Rd., Nangang Dist., Taipei City 115, Taiwan
ycshao@gmail.com

KEY WORDS: Automatic Photo Triangulation, Image Matching, Relative Orientation

ABSTRACT: In digital photogrammetry, automatic tie-point matching is necessary for photo triangulation, integrated sensor orientation, epipolar image generation, image registration, image transformation, image stitching, surface modeling, etc. The objective for this research is to find out a cost-effective way to perform automatic tie-point matching in photo triangulation. SURF (Speed-Up Robust Features) is a scale and rotation invariant feature extraction operator that is suitable for feature point matching even a certain degree of geometrical distortion exists. In this paper, based on the SURF image matching scheme, we develop three tie-point matching strategies and a robust estimation filter to remove blunder error. In the meantime, multiple images measurement is considered to increase the redundancy and internal reliability during least-squares bundle adjustment. In the experiment, the performance of the developed tie-point matching scheme is evaluated by the posterior standard error (sigma0) of image coordinate measurement after free network bundle adjustment. Images acquired from different platforms and purposes are test, such as vertical aerial DSLR images in urban area, oblique aerial DSLR images, unmanned aerial vehicle images taken in mountainous area, close-range object images. Experimental results show that the proposed scheme can match abundant of tie points and the sigma0 is less than 1/3 pixels. It demonstrates that the feasibility of the developed algorithm is high and is suitable for most of the photogrammetric applications.

1. INTRODUCTION

In digital photogrammetry, automatic tie-point matching is a prerequisite for photo triangulation, integrated sensor orientation, epipolar image generation, image registration, image transformation, image stitching, surface modeling, etc. Conventionally, in vertical aerial photo triangulation the tie-point can be matched automatically by setting up the image acquisition sequence, flight height, focal length, image size, overlap ratios, etc. However, in some cases the images were acquired from irregular viewing directions and locations, for example images taken by hand-held or rotary-wing UAV for terrestrial photogrammetric applications, the conventional auto-matching scheme will not work perfectly because the scale and geometric distortion varied a lot. A scale- and rotation-invariant image matching method is thus crucial to cope with this problem.

Lowe (2004) proposed a Scale Invariant Feature Transform (SIFT) for image matching, face detection, object recognition, etc. The proposed method include three main steps, i.e. applying DoG (Difference of Gaussian) for feature detection in scale space, feature point localization, and rotation invariant descriptors are calculated by the gradient of adjacent pixels on the detected features. Bay, et al. (2008) proposed another algorithm, called SURF (Speed-Up Robust Features), which is inspired by SIFT but is more efficient and robust to noise. It is also suitable for feature point matching even a certain degree of geometrical distortion exists.

In this paper, based on the SURF image matching scheme, we develop three automatic tie-point matching strategies and a robust estimation filter to remove blunder error. In the meantime, multiple images measurement is considered to increase the redundancy and internal reliability during least-squares bundle adjustment. In order to examine the feasibility of the proposed scheme, images acquired from different platforms and purposes are tested. For example,

vertical and oblique aerial DSLR images acquired at Tainan and Taichung City, unmanned aerial vehicle (UAV) images taken in mountainous disaster area, and close-range images for 3D surface modeling.

2. SURF

A brief introduction of SURF-based image matching scheme is described here, detail content and benchmark evaluation can be found in Bay, et al. (2008). In order to accelerate the process, SURF proposes the use of binary integral images for box type convolution filter. It represents the sum of all pixels in the input image within a rectangular region. The interest points are detected based on Hessian matrix for a blob-like structure where its determinant is the maximum. As shown in equation (1) $H(x, y, \sigma)$ is the Hessian matrix and $L_{xx}(x, y, \sigma)$ is the convolution of Gaussian second order derivative at horizontal direction with a filter size at scale of σ .

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad \text{..... (1)}$$

To improve computation efficiency, Bay, et al. adopt the approximation results of Gaussian second order derivative through integral images. However, it will introduce lower repeatability under image rotations around odd multiples of 45 degrees. Nevertheless, the result is still acceptable with high performance in computation time reduction. In order to be scale invariant, features are detected and localized on scale space. Due to the use of integral images, Bay, et al., increases filter size iteratively rather than reducing the image size to build image pyramids. The scale space consists of several octaves which represents image pyramids starts from different scale.

For each feature point, Bay, et al. utilizes the sum of Haar wavelet response to construct feature descriptor. For rotation-invariant purpose, the dominant orientation is assigned first by calculating the sum of Haar wavelet response (dx, dy) with a sliding window of 60°, where the distance is maximum. Then, it constructs a 4 x 4 square region centered at features and oriented along the dominant orientation. Finally the sum of Haar wavelet response ($\sum dx, \sum |dx|, \sum dy, \sum |dy|$) was calculated for each sub-region resulting in a descriptor vector of length 64. Feature point was matched by selecting two key points on the conjugate image with the closest Euclidian distances and their ratio (the first closest / the second closet) is less than a threshold. A smaller threshold will reduce the number of matched points but increasing its stability.

Since mismatching is unavoidable during tie-point matching, a further blunder error filtering is necessary. In the paper, based on the relative orientation concept, an iterative robust estimation scheme is designed by reducing the weight during robust least-squares adjustment (Werner, 1984). That means, when estimating the relative orientation parameters, the conjugate image coordinates (point-pair) are used to minimize the deficiency arising from these irrelevant point-pair. For ordinary least-squares adjustment, we need to minimize the weighted sum of residual squares, as shown in equation (2).

$\sum_{i=1}^{n} p_i v_i^2 = \text{Min.}$ ..... (2),

In which, n is the total number of point-pair; p_i and v_i are the weight and the residual for point-pair i, respectively. Normally, the weight is setup as 1 for all point-pairs. However, for robust-least-squares adjustment the adopted weight is defined as a function of each point-pair's residual, i.e., $p(v_i)$, which is modified according to an exponential function as shown in equation (3). The constants used are decided by empirical examinations. The weighted sum of residual squares now becomes equation (4). The conception of robust estimation is thus to reduce the weight for those irrelevant point-pair that have departed from relative orientation mathematic model. The point-pair with larger residual will contribute lesser during the iterative least-squares adjustment procedure, thus the obtained geometric model will be more accurate.

$$p(v) = \begin{cases} 1 & \text{, if } cvv > 300 \\ e^{-cvv} & \text{, if } cvv \le 300 \end{cases} \quad \text{..... (3),}$$

where $cvv = \begin{cases} 0.05 \times v^{4.4} & \text{, if No. of iteration} \le 3 \\ 0.05 \times v^{3.0} & \text{, if No. of iteration} > 3 \end{cases}$ and c=0.05.

$\sum_{i=1}^{n} p(v_i) v_i^2 = \text{Min.}$ ..... (4),

3. METHODOLOGY

Figure 1 illustrates the flowchart of the proposed photo-triangulation scheme. The input data can be the airborne vertical or oblique DSLR images, UAV DSLR images, or close-range images. According to different image acquisition methods, the matching strategies were categorized into three classes, i.e. random, sequential, and regular modes. After pair-wise image matching, each object point should have the same label on all images. The labeling is important in order to obtain redundant image measurements to the object. The tie-point matching results were imported into Photometrix Australis © for outlier removal and the estimation of exterior orientation parameters and object coordinates of the tie-point using bundle adjustment. For certain situations, we need to increase the image measurement redundancy and the strength of imaging geometry, a further back-projection point transfer by image matching is performed. That means, the object points are transferred to those unreferenced images using collinearity equations by the previous estimated exterior orientation parameters (with known interior orientation parameters) and object coordinates. Then, another SURF-based image matching is applied. In the end, another bundle adjustment is performed to finalize the photo-triangulation procedure.

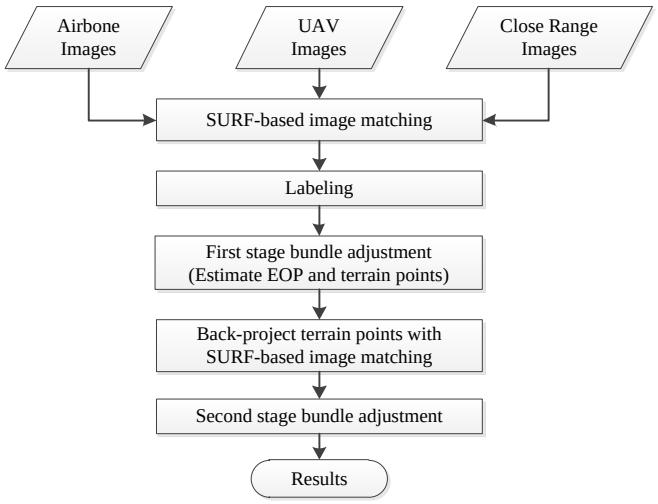


Figure 1. Flow-chart of the proposed photo-triangulation scheme.

3.1 Matching Strategies

Preventing from useless matching that introduced un-necessary computation, it is better to know the relationship between images before tie-point matching. In this study, the image acquisition method is categorized into random, sequential and regular modes. For the random mode, the camera's orientation and location are all unknown. That means, for N images there are $(N^2-N)/2$ image pairs for tie-point matching. Some of them even have no overlap, thus the matching procedure is tedious and time consuming. In sequential mode, we assume every two consecutive images have overlap on the object surface. Thus, we can perform tie-point matching for the reference one to a certain number of consecutive pictures. In case of regular image acquisition, we aware the ground coverage or geographic center of all images through the GNSS and/or IMU data. Then, the tie-point matching can be applied to those images that are close to the reference one.

3.2 Bundle Adjustment

As mentioned, the photo-triangulation is performed using Photometrix Australis© after labeling the same object points. Since the robust estimation method is applied based on the relative orientation mathematic model, some mismatching errors still exist and were eliminated manual using the Australis software as well. Thus, we address this research as "semi-automatic" photo-triangulation method. In the end, we can evaluate the tie-point matching performance by means of free network adjustment and the posterior standard error of image measurement (sigma0).

3.3 Back-projection Point Transfer

Due to the first stage of tie-point matching may not be complete; we adopt another optional back-projection with tie-point matching for point transfer from previous matching results to those unreferenced images. A pre-requisite is the exterior and interior orientation parameters have to be known in advance and the object coordinates of tie-points were estimated by the previous stage bundle adjustment. In order to equally spread the tie-points to the overlap area, we project the terrain's 3D coordinates to the unreferenced images within a 30x30 sub-regions, and then we detect the SURF features again for another tie-point matching.

4. CASE STUDY

In order to evaluate the feasibility and performance of the proposed tie-point matching and photo-triangulation scheme, six datasets that were categorized into airborne, UAV and close-range classes are tested. Table 1 illustrates the test results, including the properties of the original data, image acquisition mode, tie-point matching time, sigma0 and number of points measured from 3, 4, or 6 more images. Detail discussion will be provided in the following sections.

Table 1. Summary of test data and photo-triangulation results.

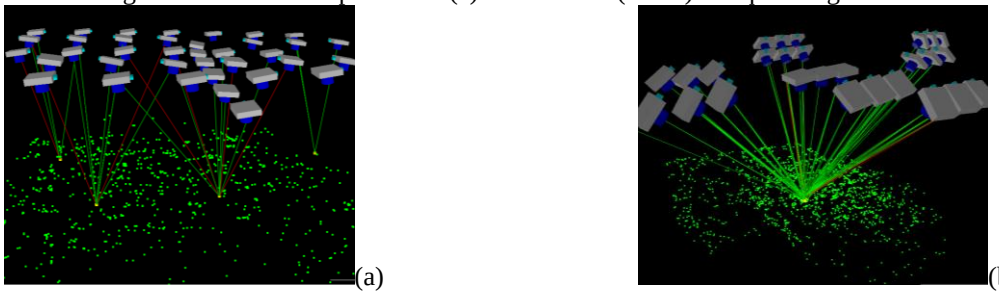
Platform	Airborne				UAV				Close Range			
Application	3D Mapping				Landslides				3D Modeling			
Case	Vertical		Oblique		Highway		Tanzi		Queen's Head		Buddha Statue	
# Images	45		27		76		156		33		10	
Camera	Sony A850		Sony A850		Sony A350		Canon		Nikon D80		Sony A900	
Image Size	6048 x 4032		6048 x 4032		4592 x 3056		4592 x 3056		3872 x 2592		6048 x 4032	
Known IOPs	Y		Y		N		Y		N		N	
Focal Length	20 mm		50 mm x 4 20 mm x 1		20 mm		24 mm		28 mm		50 mm	
Matching Strategy	Regular		Random		Random		Regular		Random		Sequential	
Matching Stage	First	Final	First	Final	First	Final	First	Final	First	Final	First	Final
Matching Time	33mn	1hr 32mn	2hr 16mn	X	3hr 32mn	X	3hr 16mn	3hr 47mn	11mn	2hr 1mn	2mn	1hr 9mn
3 points	1394	1427	1913	X	9798	X	10696	10738	2084	2104	1264	1263
4 points	304	738	82	X	3552	X	2608	4962	678	1318	451	593
6 more	34	230	3	X	622	X	389	1423	89	439	69	116
Sigma0 (pixel)	0.26	0.36	0.25	X	0.32	X	0.28	0.33	0.22	0.27	0.20	0.22

4.1 Airborne Images

The vertical and oblique DSLR images are acquired from NCKU Airborne Multiple Cameras Imaging System (AMCIS), which is designed for 3D mapping. It composes of five SONY DSLR high-resolution digital cameras, one for vertical and four for oblique imaging with tilt angle around 45 degrees. Figure 2 (a) is the vertical image; (b) & (c) are oblique ones. The correct matched points are denoted as green. As shown in Table 1, the sigma0 of both cases in the first stage are all less than 0.3 pixels, in vertical cases after back-projection with re-matching sigma0 still with 1/3 pixels. For oblique images the second stage re-matching is not test due to two different lenses are utilized. Figure 3 (a) and (b) show the 3D view of cameras' location for the vertical and oblique cases after photo-triangulation. It demonstrates several bundles to visualize the imaging geometry and the redundancy of image coordinate measurement.

(a) (b) (c)

Figure 2. Matched tie-points for (a) vertical and (b & c) oblique image cases.



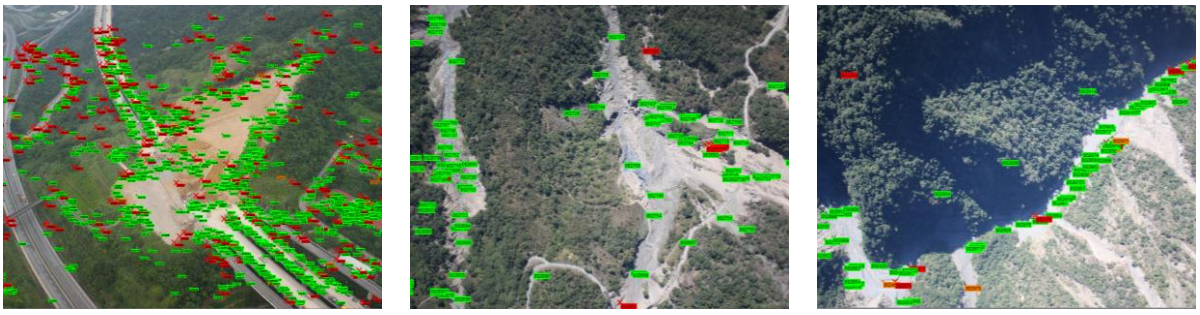
(a) (b)

Figure 3. Image distribution and point clouds of tie-points after photo-triangulation for (a) vertical and (b) oblique images.

4.2 UAV Images

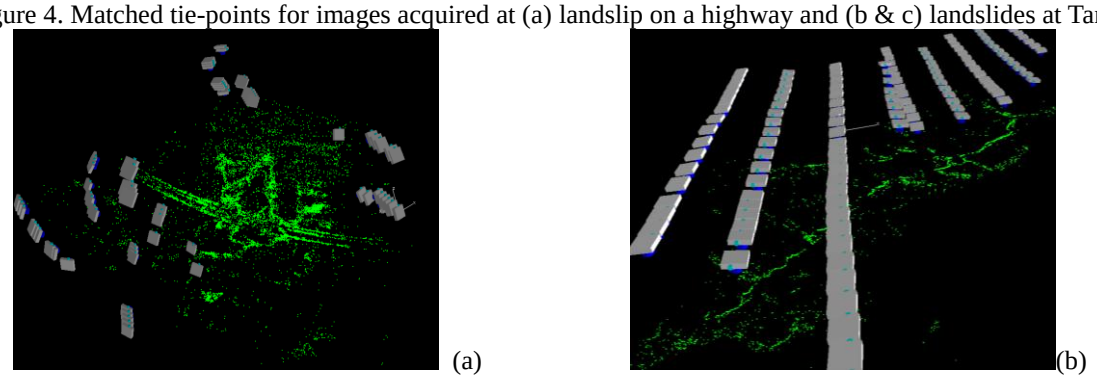
The adopted UAV images are acquired for disaster monitoring, thus a fast photo-triangulation and mapping scheme is very important. The content of the acquired UAV images are majorly landslides hazard. One is the landslip occurred at No.3 National Highway on April 25th, 2010, as shown in Figure 4 (a). The second one is landslides occurred at Tanzi, Kaohsiung City, which is induced by heavy rainfall brought by Typhon Morakot on August 8th-10th, 2009, as shown in Figure 4 (b) & (c). In this study, the matched tie-points are mostly located at roads and bare ground, except for the shadow and forestry area that may cause failures. Figures 4 (a), (b), and (c) depict the distribution of matched tie-points for both cases.

After photo-triangulation the sigma0 of highway case is larger due to the camera has not been calibrated, but still less than 1/3 pixel. Figure 5 (a) shows the distribution of images that are taken irregularly in highway case. One may clearly visualize the point clouds that belong to roads and landslip. In Figure 5 (b) we notice that the images were acquired in a systematic way. The sigma0 is less than 1/4 pixel and the point clouds for the terrain and bare ground are clear. In the meantime, since the images for landslip on a highway have high ratio of overlap, no further back-projection point transfer is applied. But the images were acquired with known GPS location for Tanzi case, the sigma0 after back-projection point transfer is still less than 1/3 pixel.



(a) (b) (c)

Figure 4. Matched tie-points for images acquired at (a) landslip on a highway and (b & c) landslides at Tanzi.



(a) (b)

Figure 5. Image distribution and point clouds of tie-points after photo-triangulation for (a) landslip on a highway and (b) landslide at Tanzi.

4.3 Close-range Images

In close-range object 3D modeling, particularly for cultural heritage digital documentation, the images may be acquired systematically or randomly. Since the cultural heritages are mostly fragile and priceless, the use of marker for automatic photo-triangulation is impossible. A markerless tie-point matching scheme is thus important to cope with the photo-triangulation problem (Barazzetti et al., 2010). In this study, two close-range cases are tested. The first one is a rock stone located at a beach and was eroded by the ocean wave for hundreds years. It was sculptured and looked like a queen's head (Figure 6). The second case is a wooden historic Buddha statue (Figure 7). The photo-triangulation result shown in Figure 8 (a) depicts that the images were acquired by surrounding the queen's head. On the other hand, as shown in Figure 8 (b), ten images were taken for the Buddha statue in a horizontal direction with convergent imaging. From Table 1 we realize that the tie-point matching scheme can also achieve both less than 1/3 pixel of sigma0 even the camera's interior orientation parameters were not calibrated in advance. In the meantime, we notice that the computation time for the point transfer took more than one hour. It is introduced by convergent imaging that all surface points are mathematically visible by all images and all of them were used in the re-matching stage for point transfer.

4.4 Discussions

From Table 1, one may also realize that the computation time for the first stage tie-point matching is highly correlated with the number of images used. However, if a-priori information for image acquisition method is provided, a suitable matching strategy can be adopted to reduce significant amount of computation time. By analyzing the number of points measured from 3, 4, or 6 more images, we found that the numbers are getting more after back-projection point transfer. It means that the image measurement redundancy and internal geometric reliability was improved by the suggested matching scheme.


Figure 6 Queen's head.


Figure 7 Buddha statue.

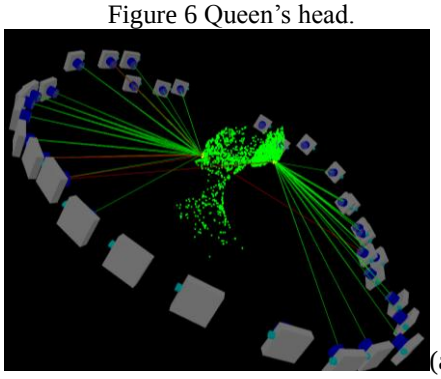
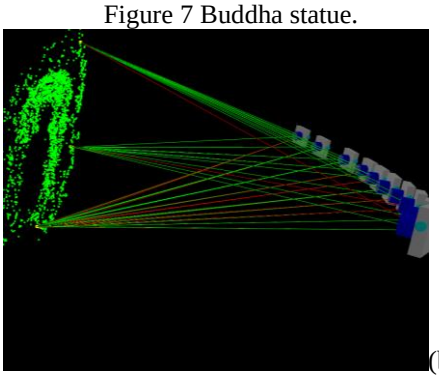

(a)

(b)

Figure 8. Image distribution and point clouds of tie-points after photo-triangulation for (a) Queen's head and (b) Buddha statue.

5. CONCLUSIONS

In this study, a SURF-based automatic tie-point matching and a semi-automatic photo-triangulation scheme is proposed and evaluated using images acquire from the airborne, UAV or hand-held close-range images. It demonstrates the feasibility of the proposed method to varied sources of frame images. It is beneficial to emergency response of natural hazard monitoring and decision making, particularly for random acquired UAV images which are difficult to perform tie-point matching using conventional vertical image matching procedure. The accuracy is evaluated through free network bundle adjustment by observing the posterior standard error of image coordinate measurement. The experimental results show that the tie-point image coordinate measurement accuracies are all less than 1/3 pixels, which is acceptable for most of the photogrammetric applications. In terms of computation time, if images were acquire in a systematic way, the amount of process time could be reduced significantly by choosing a suitable matching strategy.

Acknowledgements

This research was financially supported by the National Science Council, Taiwan (Project# NSC 100-2119-M-006-006) and Geosat Informatics & Technology Co. The authors are grateful to DPRC of NCKU and Geoforce Technologies Co., Ltd., for providing the test images.

Reference

Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., 2008. Speeded-up robust features (SURF). Computer Vision and Image Understanding, 110(3) PP. 346–359.

Barazzetti, L., Scaioni, M., Remondino, F., 2010. Orientation and 3D modeling from markless terrestrial images: combining accuracy with automation. The Photogrammetric Record 25(132) pp. 356–381.

Lowe, D.G. 2004. Distinctive image features from scale-invariant keypoints, IJCV 60 (2) pp.91–110.

Werner, H., 1984. "Automatic gross error detection by robust estimators." ISPRS Archives - Commission III, 25, Part A3, 1101-1108.