AUTOMATIC BUNDLE ADJUSTMENT OF THERMAL INFRARED IMAGES

Ling-Yi, Hsu[1] , Cheng-Yi, Chen[2], and Shih-Hong, Chio[3]
[1]Graduate student, Dept. of Land Economics, National Chengchi University,
NO.64, Sec.2, ZhiNan Rd., Wenshan District,Taipei City 11605, Taiwan; Tel: +886-2-29393091#50621;
E-mail:101257032@nccu.edu.tw

[2]Ph. D student, Dept. of Land Economics, National Chengchi University,
NO.64, Sec.2, ZhiNan Rd., Wenshan District,Taipei City 11605, Taiwan; Tel: +886-2-29393091#50621;
E-mail: 101257505@nccu.edu.tw

[3]Associate Professor, Dept. of Land Economics, National Chengchi University,
NO.64, Sec.2, ZhiNan Rd., Wenshan District, Taipei City 11605, Taiwan; Tel: +886-2-29393091#51657;
E-mail: chio0119@nccu.edu.tw

KEY WORDS: Thermal Infrared Images, Bundle Adjustment, SIFT

Abstract: Thermal infrared images show the temperature change of sensed scenes. Therefore, thermal infrared camera can sense some important information that optical images cannot do, especially in night. Thus, thermal infrared images are used not only in the domain of surveying, but also in the disaster prevention, and the environment monitoring. Before thermal infrared images can be used for advanced analysis based on the 3-D reference coordinate system, e.g. TWD97 coordinate system in Taiwan, the precise position and orientation of those images should be determined by bundle adjustment after tie points are extracted either manually or automatically. However, the manual measurement of tie points is much more cost-consuming due to the point features of thermal infrared images cannot correspond to the actual and visible point location of the scene. Therefore, in this study improved SIFT(Scale Invariant Feature Transform) Algorithm developed by Chen and Chio (2011) will be used to extract the tie points automatically, and bundle adjustment will be performed after the thermal infrared camera parameters are calibrated. This study will also discuss the relevant problems and give some suggestions from the test results.

INTRODUCTION:

There is much important spatial information which cannot be acquired from digital images. For example, thermal infrared camera can sense some important information that optical images cannot do, especially in night. In order to acquire useful spatial information based on the 3-D reference coordinate system, e.g. TWD97 coordinate system in Taiwan, from thermal infrared images, the precise position and orientation for thermal infrared images by bundle adjustment should be implemented first. Thus, some spatial information, e.g. DSM (digital surface model) and orthoimages, could be generated. While performing bundle adjustment, the keypoint is the automatic extraction and matching of tie points. Presently, the existing photogrammetric systems, e.g. ERDAS LPS, almost perform the extraction and matching by NCC (Normalized Cross Correlation). NCC algorithm is always easy to match the wrong points because it is an area-based image matching algorithm, especially, when the tie points are located in image area with low contrast and textures. In this study, because of the imaging characteristics of thermal infrared cameras, it is difficult to match correct and accurate points by NCC algorithm. Therefore, improved SIFT(Scale Invariant Feature Transform) algorithm developed by Chen and Chio (2011) will be used to extract the tie points automatically, and then the image coordinates of tie points data will be loaded into ERDAS LPS photogrammetric software to perform bundle adjustment with the calibrated camera parameters. The following sections will describe the approach to extracting tie points, the flow chart of study, test design, test results, and conclusion as well as recommendations.

IMPROVED SIFT ALGORITHM

In this study, improved SIFT algorithm, developed by Chen and Chio (2011) by including SIFT algorithm, Harris Matrix, and entropy calculation, will be used to extract the tie points automatically for matching in order to perform bundle adjustment. Therefore, this algorithm will be described more detailed as follows:

Basically, SIFT is an algorithm to extract and describe the keypoints in images. In original paper, SIFT have four main steps (Lowe, 2004):

(1). Scale-space extreme detection
(2). Keypoint localization
(3). Orientation assignment
(4). Keypoint descriptor

According to the four main steps by Lowe (2004), improved SIFT algorithm is explained below. First, SIFT creates different octaves and scales of original image. After that, Gaussian blurring is implemented on each image to generate DOG (difference of Gaussian) images, as shown in Figure 1. Secondly, many candidate points from DOG images will be acquired. Subsequently, by comparing to candidate point's 26 neighbor points, candidate keypoints will be detected, as shown in Figure 2.

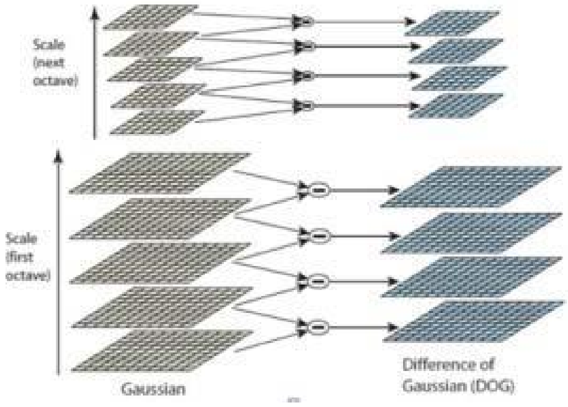

Figure 1: Difference of Gaussian (DOG) images (Lowe, 2004).

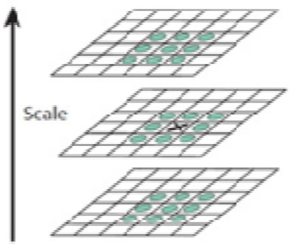

Figure 2: The 26 neighbor points are used to find the X mark whether it is maxima or minima (Lowe, 2004).

Third, because the edge keypoints or keypoints with low contrast will affect the matching results, candidate keypoints on the edge and with low contrast should be deleted. As to the points with low contrast, a threshold is set to eliminate it. For the edge points, the Harris Matrix, see equation (1), will be calculated to find them.

$$A = \sum_u \sum_v w(u,v) \begin{bmatrix} G_x^2 & G_x G_y \\ G_x G_y & G_y^2 \end{bmatrix} = \begin{bmatrix} \langle G_x^2 \rangle & \langle G_x G_y \rangle \\ \langle G_x G_y \rangle & \langle G_y^2 \rangle \end{bmatrix} \quad (1)$$

Where A is the structure tensor, $w(u,v)$ is a window over the pixel (u,v);
G_x is the gradient of x-direction, G_y is the gradient of y-direction, and angle brackets denote averaging. This matrix is called as Harris Matrix.

Besides, the rotation problems among the images should be considered while matching, the magnitude and orientation calculated by equations (2) and (3), for all pixels around the calculated keypoints will be used to resolve this problem.

$$\mathrm{m}(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (2)$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \quad (3)$$

m(x,y) is magnitude of pixel (x,y) and θ(x,y) is orientation of pixel (x,y), L(x,y) is the gray value of pixel (x,y).

For overcoming the matching problem from rotations between two images, while describing the descriptor of one keypoint which is need to decide a direction .Thus, the gradient of keypoints in all direction are calculated, then generate histogram to decide the direction of keypoints. 360 degrees for each region and calculate the sum of every magnitude in each regions to gather statistics to find the maximum value. The region with maximum is the orientation of keypoint. In this study, it uses 16x16 window to calculate histogram.

Finally, we will create a 16x16 window with a keypoint at center and calculate a histogram every 4x4 region. After all 4x4 regions' histogram done, the descriptor of this keypoint is also done, and the results are saved as a 128 dimensional vector. The vector is called the descriptor of SIFT. The descriptor should like Figure 3. By following steps above, we get keypoints from original images but it is not enough to make keypoints unique. In the texture region, it is easy to generate the similar descriptors leading to match wrong points. So we add an entropy into the descriptor to make the extected keypoints more unique.

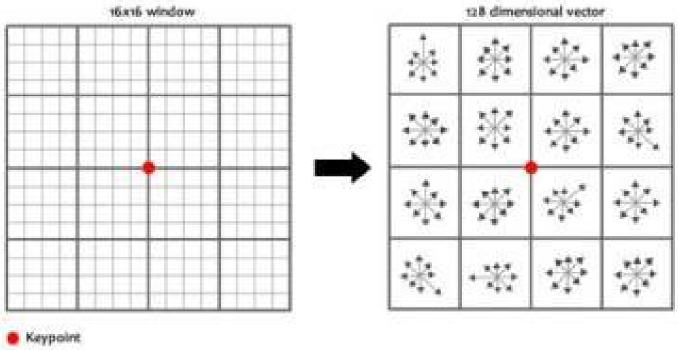

Figure 3: The illustration of candidate point descriptor (Lowe, 2004).

The entropy means a region's randomness. The equation to calculate the entropy is as follows:

$$\mathrm{e} = -\sum_{i=0}^{L-1} p(z_i) log_2 p(z_i) \quad (4)$$

Where z is the gray level values within the region, $p(z_i)$ is the number of pixels of the grayscale value z. Then by combining the SIFT descriptor and the entropy value to become a new descriptor is obtained as follows:

$$\mathrm{D} = \begin{bmatrix} \omega \times S \\ (1-\omega) \times E \end{bmatrix} \quad (5)$$

Where D is new descriptor, S is SIFT descriptor, E is an entropy value, ω is the weight, and usually ω is 0.5.

THE FLOW CHART OF STUDY

The flow chart of the study is as the following Figure 4.

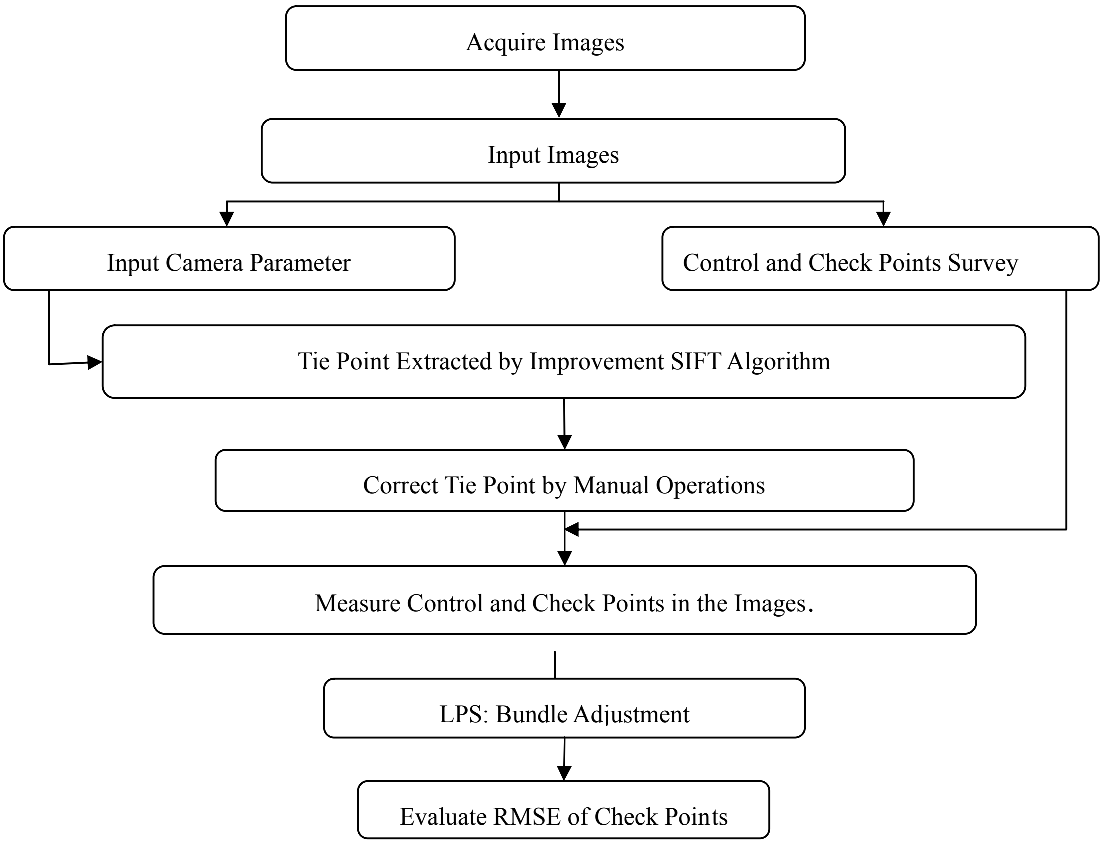

Figure 4: Flow chart of study

TEST DESIGN
The test area is a river embankment behind the college of social science in National Cheng-chi University as shown in Figure 5. The thermal infrared images are collected by thermal infrared camera FLIR T360, see Figure 6. In this study, ERDAS LPS software is used to perform bundle adjustment, and the control points and check points are surveyed by the total station. The coordinate system is TWD 97, the national coordinate system in Taiwan.

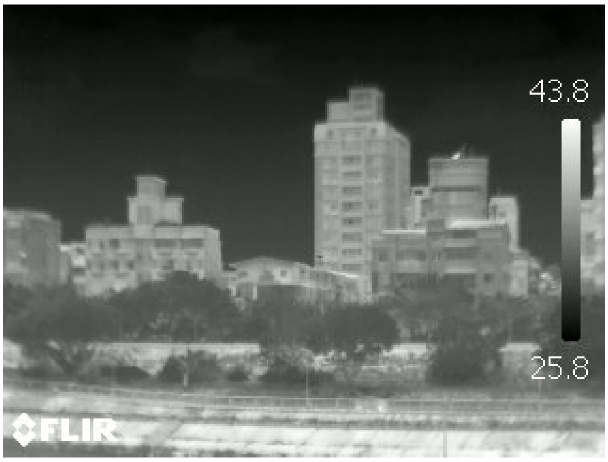

Figure 5: The image of test area

Figure 6: FLIR T360

On the river embankment, ten thermal infrared images are taken along two strips. There are five images each strip. The image end-lap is about 90%, side-lap is about 80%. The configuration of the thermal infrared images is as shown in Figure 7. For tests, the locations of taking the thermal infrared images are surveyed by a total station in order to providing initial exterior orientation for bundle adjustment.

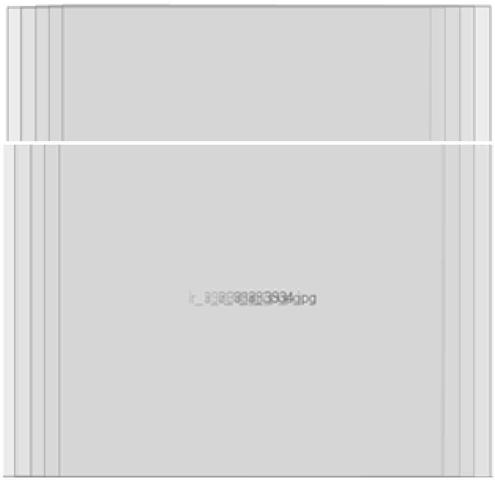

Figure 7: The configuration of images

As the flowchart shown in Figure 4, after the image are collected, improved SIFT Algorithm developed by Chen and Chio (2011) is used to extract and match the tie points automatically. Then those matching results are input into LPS software. The matching results of tie points still include some wrong matching points, so they are corrected by manual operation. Finally, the thermal infrared images, camera parameters, and control points are all loaded into LPS software for performing bundle adjustment after the control points on the images are measured. Because the matching results are verified and corrected by manual operation, no blunder detection is necessary while performing bundle adjustment.

RESULTS:

In the first test, the result of the tie points extracted and matched automatically by LPS software is analyzed. LPS uses the NCC algorithm to match the tie points. While matching, the window size is 7x7 pixels and the threshold of correlation coefficient is 0.8. There are 50-wrong-points of 163 points are generate by NCC algorithm, so the automatic matching correction rate of tie points by LPS is about 69.3% after visual inspection. Meanwhile, there are 30-wrong-points of 156 points are generate by improved SIFT algorithm, by visual inspection the automatic matching correction rate of tie points by improved SIFT algorithm is about 80.1%. According to the correction rate, we can confirm that the performance of improved SIFT algorithm is better than NCC algorithm. Therefore, in the next test we will input the tie points extracted and matched by improved SIFT algorithm into LPS software to perform the bundle adjustment and investigate the accuracy by check points. In this test, the camera parameters are calibrated by PhotoModler software. The detailed calibration can refer to the research on characteristic of area-based thermal infrared images by Na and Hung (2000). Before performing bundle adjustment, the related parameters are set, see Table 1.

Table 1: Related Parameters for Bundle Adjustment

Maximum Iteration	10
Convergence Value (meters)	0.1
Image Point Standard Deviation (pixel)	x=1,y=1
GCP Type and Standard Deviation (meters)	Same weighted values, X=Y=Z=0.05
Interior Orientation	Fixed for all images
Exterior Orientation	No weight

After the control points on the images are measured and related parameters for bundle adjustment are set, the bundle adjustment is performed. The result of bundle adjustment is shown in Table 2.

Table 2: The result of bundle adjustment

Iteration Convergence	Yes	
Total Image Unit-Weight RMSE(pixels)	0.6553	
Control Point RMSE(meters)	Ground X	0.031
	Ground Y	0.048
	Ground Z	0.0083
Check Point RMSE(meters)	Ground X	0.199
	Ground Y	0.523
	Ground Z	3.198

From Table 2, the total image unit-weight RMSE is 0.6553(pixels). The distribution of control points, check points and tie points are as Figure 8. RMSE of Check Point in three directions is 0.02, 0.52, and 3.2 meters, respectively.

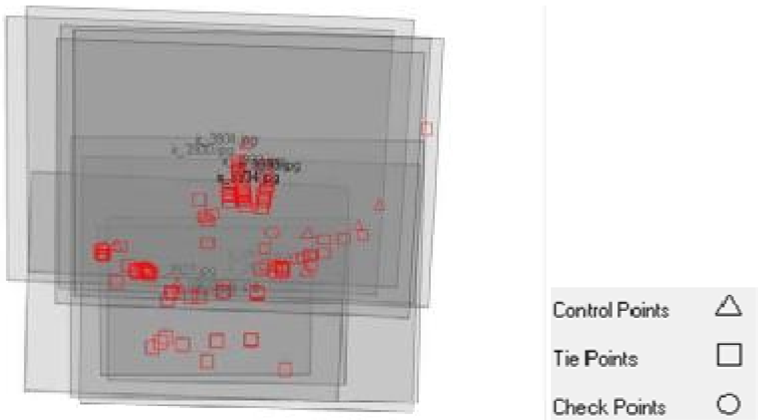

Figure 8: The distribution of control points, check points and tie points in the images

CONCLUSION AND RECOMMENTATION:

In this study the improved SIFT algorithm is used to extract and match tie points automatically. From test result, this improved SIFT algorithm can increase the matching correction rate than NCC algorithm. Besides, from the test, the resolution of thermal infrared image is too low so it is hard to recognize points that could decrease measuring accuracy and cause poor result of bundle adjustment.

Many technical issues are discovered in this study:

1. Due to the imaging characteristics of thermal infrared images, it is hard to identify and measure control points on thermal infrared images clearly and precisely by human eyes. Even though the high accuracy of control points are surveyed by total station, the results of bundle adjustment is not good enough.
2. Because the sensitivity of thermal sensor for temperature is very high, the same object may show different grayscale value in images. It will affect the result of extracting and matching tie point automatically.
3. The camera parameters are calibrated by PhotoModler program. The calibrated marks are on a flat paper without any difference of depth, which caused a poor accuracy in z-direction, i.e. depth.
4. The used image end-lap is about 90% and side-lap is about 80%. This makes the geometric distribution of images too weak .The angle of intersection is too small, that leads to the results of bundle adjustment poor. So in the future the better geometry distribution is suggested, for example the images end-lap is 80% and side-lap is 30% , to test the accuracy of bundle adjustment.
5. There are many obvious marks in the image, like thermometric scale and trademark. The tie points extracted from those marks should be excluded to avoid the wrong bundle adjustment.

The issues mentioned above will be discussed more in depth in the future.

REFERENCES:

Chih-Chung, Na;Hao-Hsiung Huang,2000.Research on characteristic of area-based thermal infrared images.

Cheng-Yi, Chen ; Shih-Hong, Chio,2010.Integrating SIFT and Harris Corner Detector for Definite Keypoint Extraction for ImageMatching.
Förstner, W; Dickscheid, T; Schindler, F, 2009. Detecting Interpretable and Accurate Scale-Invariant Keypoints In: 12th IEEE International Conference on Computer Vision (ICCV'09). Kyoto, Japan 2009, pp. 2256-2263.
Hans P. Moravec, 1980, Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover, tech. report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University & doctoral dissertation, Stanford University.
Harris, C. G. and Stephens, M. J., 1988. A combined corner and edge detector, Proceedings of the Fourth Alvey Vision Conference, Manchester, United Kingdom, pp. 147-151.
Huang, C.R., Chen, C.S., Chung, P.C., 2008, Contrast context histogram–an efficient discriminating local descriptor for object recognition and image matching. Pattern Recognition (PR 2008) 41(10), 3071–3077.
Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints, International Journal of ComputerVision, Vol. 60, No. 2, pp. 91-110.