

# Fully Automatic and Robust Approach for Remote Sensing Image Registration

Chi-Farn Chen, Min-Hsin Chen, and Hsiang-Tsu Li

Center for Space and Remote Sensing Research  
National Central University  
Jhongli, Taiwan  
cfchen@csrrs.ncu.edu.tw

**Abstract.** Image registration is an important preprocessing procedure for remote sensing image applications, such as geometric correction, change detection, and image fusion. Since it is a time-consuming and labor-intensive task to correctly register the remote sensing image, this paper proposes a fully automatic and robust approach for the remote sensing image registration. First, the image pyramid of working and reference images are constructed for coarse to fine matching processing. Second, the feature points can be automatically extracted from the reference image, and the matching point can be searched on the working image. Third, in order to improve the accuracy of registration, the robust estimation serves as an important tool in preserving the correctly matched points. Three sets of satellite images, which include multi-sensor, multi-temporal and multi-spectrum images, are used to test the proposed approach. Results show that the approach is capable of automatically registering the working image to the reference image with great precision.

**Keywords:** Image Registration, Remote Sensing Image, Automatic and Robust.

## 1 Introduction

The use of satellite images has become an increasingly important tool in environmental monitoring in recent years, as they own a large swath width, a high spatial resolution and a high receive frequency. In order to efficiently and accurately use the remote sensing images, image registration is a necessary preprocessing procedure for remote sensing image applications, such as image geo-correction, land change detection, and image fusion [1][2]. Image registration is a process that involves the establishment of a geometric mapping function between two images, which can be adopted in multi-temporal, multi-sensor, multi-spectrum and multi-resolution images. Conventionally, after the remote sensing images are acquired, the images will be corrected systematically by colinear condition models [3]. However, due to limitations in the digital terrain model accuracy and number of ground control points, various systematic geometric errors (sometimes more than 1 pixel) will exist in the satellite image products, which could result in problems when applied further. Hence, in order to improve the results of remote sensing applications, image registration is a necessary process in eliminating the geometric distortions within two images. Over

the past years, a considerable amount of studies have been made on image registration techniques and applications. The image registration method can be divided into two major categories [4], the geometric image feature based method and voxel similarity measure based method. The geometric image feature based algorithm uses the point, line or surface features to register two images via least squares fitting [5][6]. As the algorithm is a scale, offset and rotation invariant transformation, it is widely used in computer vision applications. The voxel similarity measure based algorithm maximizes the similarity between the target and search window, the most accurate registration position can be determined [7][8]. The image registration of remote sensing images is quite different to the registration on medical and computer vision images. In general, the radiometric responses of the same feature on the satellite images may vary with the viewing angles, atmospheric conditions, cloud, haze and land cover changes. Therefore, in order to perform the automatic and accurate registration of the remote sensing images, the algorithm has to take into account the radiometric variation of the features, when applied to multi-temporal, multi-sensor and multi-spectrum remote sensing images.

A fully automatic approach for accurate and robust remote sensing image registration is proposed in this paper. In light of the high efficiency advantages in the geometric image feature based method, and high accuracy in the voxel similarity measure based method; proposed approach not only uses the normalized cross correlation to accurately match the conjugate points, but also constrains the geometric relations of matching points in finding out the correctly matching ones. By repeatedly removing the points with the largest geometric errors, the incorrect matching point will be eliminated point by point. Moreover, the accuracy of the image registration can be further improved to the sub-pixel level. In order to make the image registration task more efficient and accurate, the concept of the coarse to fine matching is also applied in this research [9]. The proposed approach consists of three major steps: (1) the construction of image pyramids, (2) the detection and matching of feature points, and (3) the robust estimation and image registration. A detailed explanation of the three procedures will be presented in the following segments. The organization of this paper is as follows. Section 2 delineates the methodologies used in this research. The three different experiments that were conducted to test the proposed approach are introduced in section 3. The experimental results are shown in section 4, and the conclusions will be addressed in section 5.

## 2 Methodologies

Image registration can be regarded as a precision correction procedure, which provides the relative orientation of two images. Conventionally, the relative orientation of two images can be described by a mapping function, and the parameters of the mapping function can be calculated via finding the conjugate points on both images. In this study, an automatic feature point operator is applied to detect the feature points on the reference image and an area based matching is used to find the most similar point on the working image. By using the least squares adjustment and the coordinates of conjugate points, the parameters of the mapping function can be calculated. However, the area based matching can only provide the matched point

with the highest similarity instead of the highest correctness. Accordingly, in order to increase the accuracy of the image registration, the incorrect matched points will be continuously detected and eliminated by the robust estimation, until the registration accuracy reaches the sub-pixel level.

One of the critical processes of the proposed approach is the feature point matching. Needless to say, the efficiency and correctness of the matching process are highly dependent on the size of the search area. A larger search area will increase the computation time and decrease the correctness of the matching results. Hence, based on the coarse to fine matching concept, during the first step of the proposed method, an image pyramid is constructed on both the reference and working images. Afterwards, each image pyramid level will be registered respectively from a coarse to fine and level-by-level basis. A systematic flowchart of the proposed approach is illustrated in Fig 1. The following sections will describe the physical mechanisms of each step in greater detail.

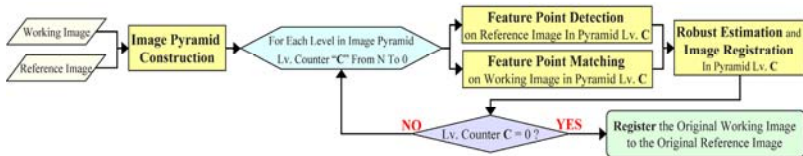
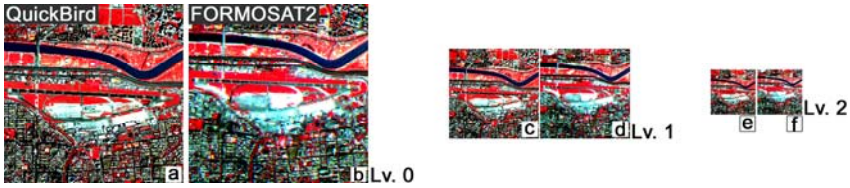


Fig. 1. Flowchart of the proposed approach

## 2.1 The Construction of Image Pyramid

The purpose of this paper is to register the working image by finding the correct matched feature points in both the working and reference images. However, due to influences from the viewing angle, atmospheric conditions and land cover changes, it is impossible for two temporal images to be exactly the same. Accordingly, the image pyramid concept is employed in the first step. The image pyramid is a hierarchical structure and the original image is the Lv.0 of pyramid. Lv.0 is low-pass filtered and sub-sampled by a factor of two to construct the next pyramid level. Further repetitions of the filter/sub-sample steps generate the remaining pyramid levels. The low-pass-filter properties of the image pyramid not only can reduce the differences between the two images but also increase the efficiency and accuracy of the matching task. Fig. 2 is a simple example in illustrating the advantages of the image pyramid construction procedure. A QuickBird image with a higher spatial resolution (2.4m) is regarded as the reference image, while a FORMOSAT2 image with a lower spatial resolution (8m) is adopted as the working image. Results of the image pyramids of both the QuickBird and FORMOSAT2 images are presented respectively in Fig. 2. Fig. 2(a) and (b) are the original images at the first level of the image pyramid; (c) and (d) are both the images at the second level, (e) and (f) are the images at the third level. Since the image pyramid can improve the accuracy of the matching results, and also shrink the boundary of the search area, its construction is a necessary pre-processing step for the following procedures.



**Fig. 2.** Illustration of image pyramids. (a) and (b) are original images of QuickBird and FORMOSAT2 at the first level (Lv. 0). (c) and (d) are both images at the second level (Lv.1). (e) and (f) are both images at the third level (Lv.2).

### 2.2 Detection and Matching of Feature Points

In order to establish the mapping function between the reference and working images in each image pyramid, it is required to find the conjugate points on both images. Hence, the proposed method detects the feature points on the reference image, and searches the matching points on the working image. It is clear that the large amount of well-distributed conjugate points can improve the accuracy of the image registration. The Target Defined Ground Operator (TDGO) is employed in this approach to automatically and uniformly detect a significant number of feature points from the reference image [11]. A bit pattern table is used in TDGO to detect the points with the right-angle features. The bit patterns of feature points are illustrated in Fig. 3(a). In order to quantitatively measure the pattern of a 3 x 3 window, a bit value (BV) calculated from the grey value differences in a 3 x 3 window is defined in Eq. (1) and (2). Therefore, the points that have exactly the same BV's as the patterns of feature points (Fig. 3(a)) will be regarded as the feature points. The advantage of TDGO is the ability of identifying the points with right-angle features, which is critical for image matching. Accordingly, an area-based matching algorithm for each feature point is subsequently applied to find the matching points on the working image. This study selects the Normalized Cross Correlation (NCC) as the objective function for the area-based matching. The NCC measures the correlation via a two blocks' comparison. A higher value of the NCC means a higher probability of the two blocks being similar. Since the NCC is a normalized operation, the matching results will not be influenced by illumination conditions. Thus, the NCC is more suitable for the application of satellite image matching.

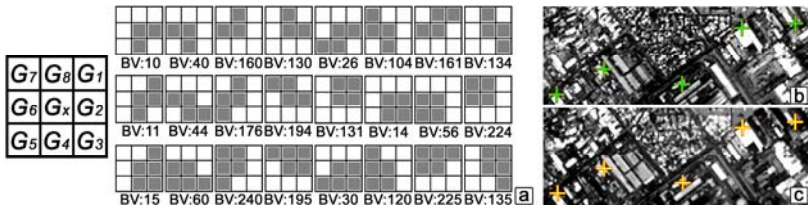
$$if (|G_x - G_i| \geq Threshold) B_i = 2^{i-1} \text{ else } B_i = 0 \quad i = 1, 2, \dots, 8 \tag{1}$$

$$BV_x = \sum_{i=1}^8 B_i \tag{2}$$

Where  $G_x$  is the grey value at pixel  $x$ ;  $G_i$  are the grey value of 8 neighbouring pixels;  $B_i$  is the binary code transformed from the grey level differences; the TDGO *Threshold* used in this study is 70 and the  $BV_x$  is the Bit Value at pixel  $x$ .

After the NCC matching procedure, each feature point, which is detected by TDGO on the reference image, can find a matching point with the highest correlation coefficient on the working image. The value of the correlation coefficient is between  $-1$  to  $1$ , where a higher correlation coefficient value signals a higher similarity of two matching blocks. In this study, 0.75 is used as a threshold to evaluate

the matching results. It is deemed successful when the correlation coefficient value is higher than the threshold, and vice versa when it is lower. Fig.3 indicates the results of the feature point detection and matching procedure. The green crosses in Fig. 3(b) denote the feature points, which are detected automatically by TDGO on the reference image pyramid Lv. 1(QuickBird), while the yellow crosses in Fig. 3(c) represent the success matching points on the working image pyramid Lv. 1(FORMOSAT2). The area-based matching technique is a viable tool for finding success matching points. However, if there are several similar features on the searching area (e.g. the corner of building or the central line of the runway), it may lead to incorrect matching results. Consequently, the next logical step is to preserve the correct matching points from the success matching points, and register the working image via the construction of a mapping function.



**Fig. 3.** Results of the feature points detection and matching. (a) Patterns of feature points used in TDGO. (b) Feature points (green crosses) detected on the reference image. (b) Success matching points (yellow crosses) on the working image.

### 2.3 Robust Estimation and Image Registration

The detection and feature points matching can automatically extract the success matching points on the working image. However, it is difficult to guarantee that the point with the highest similarity will be the correctly matched one. Hence, it is necessary to remove the incorrect matching points, as they decrease the accuracy of the image registration. Therefore, based on the assumption that all of the conjugate points have common transformation parameters, the robust estimation is a powerful criterion in preserving the correct matching points. Since both the working and reference images have been systematically corrected, the assumption used in this step is reasonable. The geometric relation between the working and reference image can be described by the Affine-Transformation in Eq. (3). By using the least squares adjustment, the variation of the observation equation is an important index that notably figures out which point has the largest bias relative to the others. The equations are shown in Eq. (4).

$$X = ax + by + c \quad Y = dx + ey + f \tag{3}$$

$$\begin{bmatrix} v_{1x} \\ v_{1y} \\ v_{2x} \\ v_{2y} \\ \vdots \\ v_{nx} \\ v_{ny} \end{bmatrix}_{[2n][1]} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_n & y_n & 1 \end{bmatrix}_{[2n][6]} \cdot \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}_{[6][1]} - \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ X_n \\ Y_n \end{bmatrix}_{[2n][1]} \tag{4}$$

Where:

- $n$  : Number of success matching points
- $(X_i, Y_i)$  : Coordinates of feature point  $i$  in reference image
- $(x_i, y_i)$  : Coordinates of matching point  $i$  in working image
- $(v_{ix}, v_{iy})$  : Transformation variations of point  $i$  in  $x$  and  $y$  axis
- $[a \ b \ c \ d \ e \ f]^T$  : Parameters of Affine Transformation

By calculating the variation of each success matching point, the point  $P_L$  with the largest variation can be found. If the variation of  $P_L$  is larger than the given threshold, it denotes that the accuracy of the matching point  $P_L$  is not good enough, where it will be eliminated from the success matching points. The removal of the  $P_L$  and the recalculation of the preserved point variations will proceed repeatedly until the largest variation is smaller than the threshold. However, when the working image is totally different with the reference ones, and the number of success matching points is less than 10 points, the robust estimation procedure will stop, resulting in the image registration task to fail. If the procedure of the robust estimation is completed, the working image can be correctly registered to the reference image via the parameters of the Affine Transform. The flowchart of the robust estimation and image registration is illustrated in Fig 4.

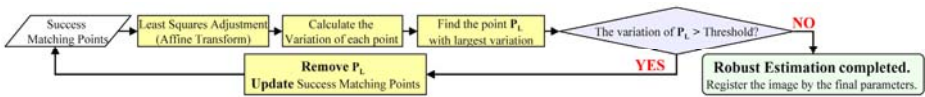


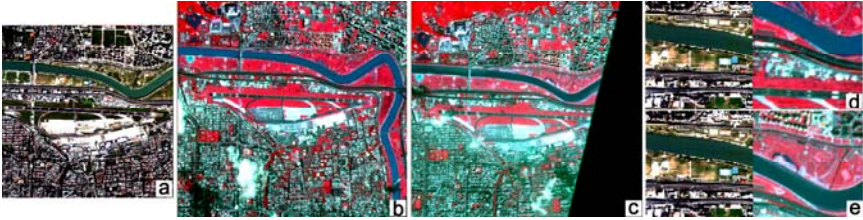
Fig. 4. Flowchart of robust estimation and image registration

As mentioned above, the proposed method is a coarse to fine matching approach, and the image registration task is processed by a level-by-level and coarse-to-fine procedure. Hence, the highest level of the working image pyramid will be registered to the referenced ones during the beginning. Since the working image has already been refined in the previous level, the size of the searching area used in the feature points detection and matching step can be decreased. Moreover, the efficiency and accuracy of the image registration can be improved level-by-level.

### 3 Experiments

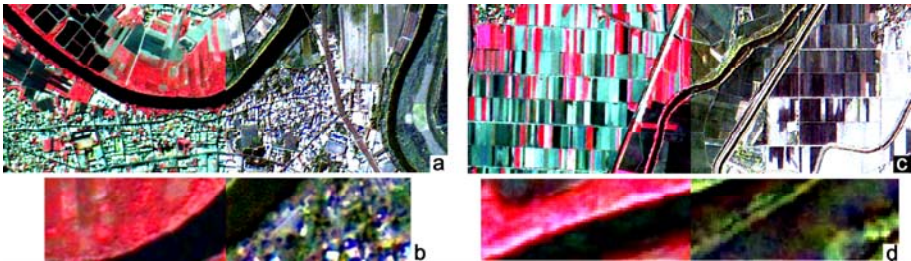
Three different experiments containing the “Multi-sensor”, “Multi-temporal” and “Multi-spectrum” images are used to test the practicability and robustness of the proposed approach. The purpose of the “Multi-Sensor” experiment is to test the ability of the proposed approach in robustly correcting the satellite image. A roughly corrected satellite image can register to a well-corrected one by the image registration technique. Therefore, in the “Multi-sensor” experiment, the reference image is a well geometric corrected QuickBird satellite image, where the accuracy of the image is less than 2 meters. The working images used in this experiment are two poor quality corrected FORMOSAT2 images, which are corrected only by the satellite orbit parameters. The error in the image can reach up to 300 meters. Moreover, in order to

perform the robustness test of the proposed method, two images containing clouds and featuring a distinctive black border are selected as the working images. In order to distinguish the reference image from the working one, it is shown in true color in Fig. 5(a), while the two working images are depicted by a false color in Fig. 5(b) and 5(c). The ground coverage of the reference and working images are 3km by 3km and 4km by 4km, respectively. The geometric difference between the reference and working images can be observed clearly in Fig. 5(d) and 5(e) after they are overlapped.



**Fig. 5.** Test images used in the “Multi-sensor” experiment. (a) Reference image. (b) Working image 1. (c) Working image 2. (d) and (e) are image enlargements of the overlapping reference and working images.

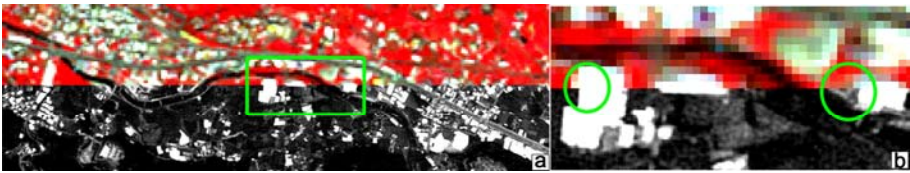
Land change detections are an important application of remote sensing images. However, to ensure a successful change detection, there should be no relative distortion between the temporal images. Since it is difficult to satisfy the condition merely through a systematic geometric correction, the image registration technique serves as a powerful tool in eliminating the relative distortion of the temporal images. Two different sets of images are tested in the “Multi-temporal” experiment. Both the reference and working images are well geo-corrected FORMOSAT2 images and taken respectively on 02/06/2007 and 02/27/2007. The reference images are shown in false color in the left of Fig. 6(a) and 6(c), and the working images are illustrated by a true color in the right. Since both test images are well geo-corrected, there appears to be no obvious geometric gap when the two images are overlapped. However, the small geometric gaps (about 10 meters) can still be observed in both two sets from the border between the two images in Fig. 6(b) and 6(d).



**Fig. 6.** Test images used in the “Multi-temporal” experiment. (a) The overlapping of the reference and working image 1. (c) The overlapping of the reference and working image 2. (b) and (d) are image enlargements of the overlapping reference and working images.



The “Multi-Spectrum” experiment is designed for image fusion applications. Caused by the different spatial resolution in the multi-spectrum (8m in FORMOSAT2) and panchromatic (2m in FORMOSAT2) images, they undergo the systematic geometric correction by their own respectively. Accordingly, even though the multi-spectrum and panchromatic images were taken within the same time frame, small geometric displacements may still occur. In this experiment, the FORMOSAT2 panchromatic image, which owns a higher accuracy of the image ground position, serves as the reference image. The FORMOSAT2 multi-spectrum image, on the other hand, functions as the working image. The overlapping of the two different images is shown in Fig. 7(a), where an image enlargement is shown in Fig. 7(b). Some displacements are seen to exist between the reference and working image. (Marked by the green circle in the image enlargement).



**Fig. 7.** Test images used in the “Multi-spectrum” experiment. (a) The overlapping of the reference (Panchromatic 2m resolution in bottom) and working (Multi-spectrum 8m resolution in top) images. (b) Image enlargements of the overlapping reference and working images.

## 4 Results and Discussions

The experimental results in this research were evaluated by overlapping the reference and registered images for visual comparisons, along with calculations of the Root Mean Square Error (RMSE) of the check points in the registration accuracy evaluation. The visual comparison results of the three experiments (Multi-sensor, Multi-temporal and Multi-spectrum) are shown respectively in Fig. 8, Fig. 9 and Fig. 10; the RMSE of the check points are presented in Table 1. Two working images were tested in “Multi-sensor” experiment via the overlapping of the reference and working images. Fig. 8(a) and Fig. 8(b) both demonstrate that the proposed approach can correctly register the Working Image 1 and 2 to the reference ones. Table 1 also indicates that the RMSE of the check points have been substantially decreased to a level less than 2 pixels. Even though the black border and clouds cover resulted in differences between the working and reference images, the proposed approach can still correctly register them. As that the working images used in the “Multi-sensor” experiment were geo-corrected only by the orbit parameters, it is impossible to correct the geometric distortion without the terrain model. Consequently, the roughly geo-corrected working images in the “Multi-sensor” case led to a higher RMSE in the check points.

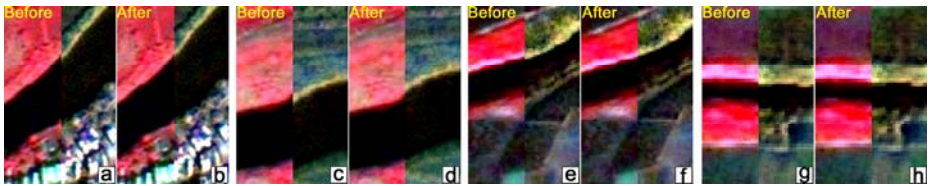
In the “Multi-temporal” experiment, since both the reference and working image are systematically geo-corrected, the geometric displacement between the temporal images is relatively small (Fig. 9(a)(c)(e) and (g)). However, the proposed approach can still improve the accuracy of the working image to the sub-pixel level (Fig. 9(b)



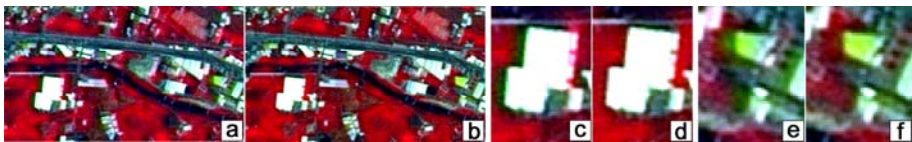
(d)(f) and (h)). Fig.10 indicates the difference of the fusion results between, before, and after the image registration process in the “Multi-spectrum” experiment. The color-mismatch phenomenon can be clearly observed in the fusion results without the image registration preprocessing (Fig. 10(a) (c) and (e)). The fusion results with the image registration preprocessing are shown in Fig. 10 (b) (d) and (f). As both the figures and the table indicate, the proposed approach can successfully and accurately be applied to the multi-spectrum applications.



**Fig. 8.** Overlapping of the reference and registered working images in the “Multi-sensor” case. (a) The registration results of Image 1. (b) The registration results of Image 2.



**Fig. 9.** Overlapping of the reference and working image before and after registration in the “Multi-temporal” case. The image enlargements of the overlapping before the registration are shown in (a) (c) (e) and (g); the results after the registration are shown in (b) (d) (f) and (h).



**Fig. 10.** The fusion results before and after the image registration process in the “Multi-spectrum” case. The fusion results without the image registration are shown in (a), (c) and (e); the fusion result with the image registration preprocessing are shown in (b), (d) and (f).

**Table 1.** The Root Mean Square Error of the check points before and after the registration

Case	RMSE of Check points (Pixels) <i>Before</i> Registration	RMSE of Check points (Pixels) <i>After</i> Registration
Multi-Sensor (Image 1/2)	38.9 / 7.1	1.7 / 1.9
Multi-Temporal (Image 1/2)	3.2 / 1.9	0.9 / 0.5
Multi-Spectrum	1.1	0.6

## 5 Conclusions

This paper introduces a fully automatic approach for an accurate and robust remote sensing image registration. The proposed approach consists of three major steps – Step1: Construction of the image pyramids of both the reference and working image. Step2: In order to establish the geometric relation between the two images, the feature point detection and matching techniques are applied to find the conjugate points. Step3: The working image can be accurately registered to the reference image by the robust estimation. Three different sets of satellite image were designed to test the reliability of the proposed approach. The experimental results indicate it can indeed successfully and accurately register the working image to the reference image. This hold true even if the two images were taken by different sensors or varying time frames and spectrums. These results prove that the proposed approach can successfully be applied as a preprocessing procedure to the remote sensing image application, such as geometric correction, land change detection and image fusion.

## References

1. Li, W., Leung, H.: A Maximum Likelihood Approach for Image Registration Using Control Point And Intensity. *IEEE Transactions on image processing* 13(8), 1115–1127 (2004)
2. Mao, Z., Pan, D., Huang, H., Huang, W.: Automatic registration of SeaWiFS and AVHRR imagery. *Int. J. Remote Sensing* 22(9), 1725–1735 (2001)
3. Chen, L.C., Teo, T.A., Rau, J.Y.: Optimized patch back-projection in ortho-rectification for high resolution satellite images. In: *IAPRS*, pp. 586–591 (2004)
4. Mäkelä, T., Clarysse, P., Sipilä, O., Pauna, N., Pham, Q.C., Katila, T., Magnin, I.E.: A Review of Cardiac Image Registration Methods. *IEEE Transactions on medical image* 21, 1011–1021 (2002)
5. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-D point sets. *IEEE Transactions on Pattern Anal. Machine Intell.* PAMI-9, 698–700 (1987)
6. Umeyama, S.: Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Anal. Machine Intell.* 13, 376–380 (1991)
7. Van Den Elsen, P.A., Pol, E.D., Sumanaweera, T.S., Her, P.F., Napel, S., Adler, J.R.: Grey value correlation techniques used for automatic matching of CT and MR brain and spine images. In: *Proc. SPIE Visualization in Biomedical Computing*, vol. 2357 pp. 227–237 (1994)
8. Netanyahu, N., Le Moigne, J., Masek, J.: Geo-Registration of Landsat Data by Robust Matching of Wavelet Features. *IEEE Transactions on Geoscience and Remote Sensing* 42(7), 1586–1600 (2004)
9. Stone, H.S., Le Moigne, J., McGuire, M.: Image Registration Using Wavelet Techniques. In: *Proceedings of SPIE*, vol. 3240, pp. 116–125 (1998)
10. Thevenaz, P., Ruttimann, U.E., Unser, M.: A pyramid approach to sub-pixel registration based on intensity. *IEEE Transactions on image processing* 7, 27–41 (1998)
11. Chen, L.C., Lee, L.H.: Progressive Generation of Control Frameworks for Image Registration. *Photogrammetric Engineering and Remote Sensing* 58(9), 1321–1328 (1992)