An Accurate Feature Point Matching Algorithm for Automatic Remote Sensing Image Registration

Guan-Long Wu and Herng-Hua Chang*
Computational Biomedical Engineering Laboratory
Department of Engineering Science and Ocean Engineering
National Taiwan University, Taipei, Taiwan
lion811004@gmail.com, herbertchang@ntu.edu.tw

***Abstract*—Remote sensing image registration is still a challenging task because of diverse image types and the lack of a consistent transformation. To improve image registration in remote sensing, this paper develops a robust and accurate feature point matching framework. A modified scale-invariant feature transform (SIFT) method is first introduced for feature detection and pair matching. Based on the properties of matched pairs, the standard grouping boundary (SGB) and confidence elliptical boundary (CEB) are computed for further examination. The SGB is utilized to categorize matched pairs according to the pair slopes. The CEB is further employed to remove outliers whose feature positions are outside the elliptical contour. Finally, the random sample consensus (RANSAC) approach is implemented to acquire the most appropriate transformation function for image registration. The proposed image registration algorithm has been tested on numerous multi-temporal remote sensing images. Experimental results validated the improvement in feature matching accuracy, which resulted in better registration performance over state-of-the-art methods. This new framework is of potential in many remote sensing applications that require automatic image registration.**

Index Terms—image registration, scale-invariant feature transform, outlier removal, remote sensing image

I. Introduction

Remote sensing images play an important role in recent environmental studies such as crop yield observation, water resources of reservoirs, the impact of human activity, and the rising of the sea level [1]. A key to the abovementioned issues is an efficient image registration method that can automatically perform accurate transformation. For remote sensing, image registration can be described as a process of matching followed by aligning two or more images taken from the same scene at different times, by different sensors, or from different views. The main concept of automatic image registration is to find an accurate set of control points and obtain the transformation matrix, which is most appropriate to the pair of images being registered.

Generally speaking, automatic image registration algorithms include three steps [2]: a) extraction of distinct regions, b) matching of the feature by searching for a transformation that best aligns them, and c) resampling one image to construct a new image in the coordinate system of the target image. One popular technique has been the scale-invariant feature transform (SIFT) [3], which includes the extraction of features that are invariant to the image scale and rotation as well as robust matching, which could endure a substantial range of noise and distortion. In the field of remote sensing image registration, SIFT has been utilized to find the invariant control points such as corners of crossroads and to match the pairs of control points between the reference image and the sensed image. However, the resembling objects may lead to errors in the matching process. To reduce the mismatched pairs of control points, Li et al. [4] proposed a robust SIFT framework to enhance the correct-match rate by feature descriptor refinement. Gong et al. [5] and Li and Ye [6] organized a SIFT control point outlier removal process to take away mismatched pairs. Though many methods [7-10] have been proposed, automatic registration for remote sensing still remains a challenge. In this paper, we develop a new scheme to refine the control points in the process of the SIFT detector and introduce a robust outlier removal approach to dispose mismatches. The outcome is a more reliable and accurate image registration algorithm for remote sensing.

II. Background

A. Image Registration

In remote sensing, image registration is the process of mapping a sensed image to another image with a transformation function computed by two sets of control points. For example, let $I_1(x, y)$ and $I_2(x, y)$ respectively represent a sensed image and a registered image, which are given 2-D matrices of images and (x, y) is defined as the discrete coordinates. We map the sensed image to the registered image by the following expression:

$$I_2(x, y) = h(I_1(T(x, y))) \quad (1)$$

where T is a transformation function and h is a resampling function. The transformation function is associated with the corresponding control point pairs $\{(x_i, y_i), (x_i', y_i')\}$ with $i = 1,2,3 \dots n$. The relationship between the control point pairs and the transformation function can be expressed as

$$(x_i', y_i') = T(x_i, y_i) \quad (2)$$

The parameters of the transformation function can be obtained by estimators. When mapping the pixels of the sensed image, a

978-1-5386-2839-3/17/$31.00 ©2017 IEEE

resampling process is required to solve the problem of the non-integer mapped coordinates.

B. Point Detector and Descriptor

Point detectors are utilized for feature point identification. The most common method is to find the local extrema in intensity in the spatial domain and the locations are considered the feature points with critical information. Based on the Hessian matrix, a wide variety of point detectors have been proposed. The method proposed by Beaudet [11] detects feature points based on the second derivative of image magnitudes. The author employed the Hessian matrix to find feature locations, where the absolute value of the Hessian matrix determinant is larger than a prescribed threshold, which becomes a local maximum. Subsequently, Harris and Stephens [12] developed an approach for detecting corners using the eigenvalues of the Hessian matrix.

Alternatively, the Laplacian of Gaussian (LoG) is a classical method to compute the second derivative by convolving the image with a Gaussian kernel function for smoothing. The Gaussian kernel function is treated as a pre-processing step to remove noise. The local extrema of the LoG then indicate the centers of bright or dark blobs in an image. The difference of Gaussian (DoG), which was suggested by Marr and Hildreth [13], is an approximation of the LoG, which was adopted by Lowe [14] to identify feature locations. In SIFT, the location is found if the value is a local extremum in a three-dimensional scale-space. In remote sensing image registration, point detectors are critical to the success of finding the feature points that are invariant to the time-scale.

A feature descriptor carries information regarding the local features in an image. The information could be a value or a vector and an image descriptor is utilized to distinguish one local feature from another. In general, descriptors are vectors obtained by the components of statistical and spatial properties of an image. To create a descriptor, a window centered at the feature points is needed for counting and classifying the features inside the window. One class of feature descriptors exhibits invariant to rotation such as histogram-based descriptors. Another class calculates the direction of the gradient at a key point as the dominant direction and computes the relative properties based on the known orientation. After the process of computing the descriptors of feature points, two sets of feature points can be matched by the least distance of descriptors and a threshold called “distance ratio”.

C. Mutual Information

Mutual information (MI) is a common measure of the amount of information that one variable contains about another. Russakoff et al. [15] used the MI as a similarity measure in images. The MI has been adopted in many clinical applications, e.g., Pluim et al. [16] introduced the medical image registration based on the MI. In the field of remote sensing image registration, the MI is a similarity measure to compare transformed images with reference images. Shannon [17] first proposed a logarithm measure based on the entropy as follows

$$H = -K\sum_{i=1}^{n} p_i \log(p_i) \tag{3}$$

where K is a positive constant for the choice of measures and p_i is a probability function. The reason for using the logarithmic measure is that the engineering parameters such as time and bandwidth have a tendency to vary linearly with the logarithm of the number of possibilities, which is more intuitive.

Let X and Y be random variables. The entropy of these two variables can be represented as:

$$H(X,Y) = -\sum_y \sum_x p(x,y)\log(p(x,y)) \tag{4}$$

The marginal entropy is defined as:

$$H(X) = -\sum_i^n p(x_i)\log\big(p(x_i)\big) \tag{5}$$

For two images A and B, the MI is defined as:

$$MI(A,B) = H(A) + H(B) - H(A,B) \tag{6}$$

There are four basic properties of the MI [16]:

1) $MI(A,B) = MI(B,A)$
2) $MI(A,A) = H(A)$
3) $MI(A,B) \le H(A), MI(A,B) \le H(B)$
4) $MI(A,B) \ge 0$

The MI is associated with one image corresponding to regions in the other image that has similar grey values.

D. Root Mean Square Error

Root mean square error (RMSE) is frequently described as an evaluation criterion for remote sensing image registration [4-6, 10]. Especially for multi-temporal remote sensing images, two images are basically different but invariant features can still be found. Based on invariant features, two images are aligned and observed by the difference between the regions. Therefore, the matching accuracy of multi-temporal remote sensing image registration is evaluated by the RMSE instead of the MI. The MI is more appropriate for two images with the same distribution of feature intensities.

A number of corresponding key point pairs are chosen manually from the sensed image and the reference image, which is viewed as the ground truth. Consequently, those key point pairs are utilized as the reference to examine the precision of the model parameters. The RMSE is mathematically defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}[(x_i - x_i'')^2 + [(y_i - y_i'')^2]} \tag{7}$$

where (x_i'', y_i'') is the transformed coordinates of (x_i', y_i').

III. Related work

A. SIFT

SIFT proposed by Lowe [3] is a robust method for finding and matching the feature points with the most similar local descriptors that consists of seven steps:

1) Gaussian smoothing: Different Gaussian filters with gradual increasing scale levels are employed to produce stacks of images. A stack of images is called an ‘octave’. The Gaussian filter, $\mathrm{G}(x,y,\sigma)$, is defined as the two-dimensional and variable-scale Gaussian kernel. The smoothed image $\mathrm{L}(x,y,\sigma)$ is defined as the convolution of the input image $\mathrm{I}(x,y)$ with $\mathrm{G}(x,y,\sigma)$. In each octave, the basal-level image is repeatedly convolved with the Gaussian filters with increasing scale levels of σ. After the basal octave, the base-level image is down-sampled by taking every other pixel in each row and column.
2) DoG: To simulate the LoG effect, the DoG is considered the approximate method. From the viewpoint of mathematics, LoG is regarded as the partial derivative of

the DoG with respect to the scale level σ. The DoG is obtained by subtracting adjacent images with different scale levels in each octave.

3) Local extremum detection: To find the local extrema of the DoG, each point is examined by adjacent points including its eight neighbors and nine neighbors below and above the current scale. That is to say that points located on the edge of three-dimensional octaves are not considered. However, in the case of remote sensing hundreds of feature points even thousands are found, so the process of refinement is essential to remove points that are sensitive to noise and located along an edge.
4) Key point refinement: Although point candidates in each octave are found and regarded as locations of key points, an additional method is used to make sure that the values at candidate locations are also the extrema in the scale-space. Another approach [18] has been proposed to find the maxima and minima in a Laplacian pyramid in the scale-space, which can be applied to SIFT for refinement because the DoG is the approximation of the LoG. Besides, removing points with a strong response along edges can be based on two eigenvalues of the Hessian matrix.
5) Orientation assignment: To create a rotation-invariant image descriptor, the information of descriptors can be related to the dominant direction computed by image gradients and feature points. Therefore, images can be matched based on the most similar feature information between two images. An orientation is assigned by a weighted histogram of gradients within a circular window with a certain radius. The direction is determined according to the strongest peak in the histogram. Gradients at sample points inside the window are computed using:

$$m(x,y) = \sqrt{\big(L(x+1,y)-L(x-1,y)\big)^2 + (L(x,y+1)-L(x,y-1))^2} \quad (8)$$

$$\theta = tan^{-1}\left(\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)}\right) \quad (9)$$

where $L(x,y)$ is the Gaussian smoothed image in the scale-space, $m(x,y)$ is the gradient magnitude, and θ is the gradient direction.

6) Local descriptor: Based on the property of the gradient, the magnitude and direction are used to describe feature points in a local region. After knowing the dominant direction of the feature points with respect to the image coordinates, the descriptors can be rotated afterward according to the dominant direction to achieve rotation-invariance. The information is then weighted by a Gaussian function with σ equal to one half of the square window width. The purpose of using the Gaussian function is to reduce the effect of sample points that are far from the center and to achieve rotation-invariant by a circular window. A key point descriptor is subsequently created by these sample gradient magnitudes, which are accumulated in eight directions and normalized in each window. A 4×4 window is adopted as a descriptor vector that contains 128 values of the normalized gradient magnitudes.
7) Matching: This process is to find the minimum Euclidean distance of descriptor vectors between the input image and the reference image. The minimum distance indicates that the key feature is most similar to other features. However, this could lead to matching errors. The second minimum Euclidean distance can be a good estimate for the false matches. Consequently, a measure called '*distance ratio*' is proposed as a threshold for selecting appropriate features.

B. Random Sample Consensus

The random sample consensus (RANSAC) method [19] is commonly adopted to compute the fittest model for some set of data. RANSAC starts from choosing a subset of data as samples to pre-compute the sample model. Then, the system counts how many data are involved in the sample model within a small tolerance and the involved data are defined as the consensus set. The abovementioned steps are repeated several times in order to find the largest consensus set. RANSAC has been extensively applied to the fields of image registration and image stitching for removing mismatches and computing the most appropriate transformation function.

C. Outlier Removal

Incorrect correspondences often have a negative effect on the accuracy of image registration, even just one pair of wrong matches. As such, a process to remove mismatched feature points is necessary. Measuring the distance ratio in the SIFT process is one of the approaches to remove outliers. The equation of the distance ratio can be expressed as:

$$d^{st} < d^{nd} \times dR \quad (10)$$

where d^{st} is the smallest distance between descriptors, d^{nd} is the second smallest distance, and dR is the distance ratio within [0 1]. If dR is low, it indicates that d^{st} is much smaller than d^{nd}. Although the distance ratio is a robust method to remove mismatches, many correct matches are removed when d^{st} is too close to d^{nd}.

To increase the accuracy of image registration, Gong et al. [5] proposed a modified outlier removal method, which defined the Gong ratio GR_{ij} as

$$GR_{ij} = \frac{R_{ij}}{S_{ij}} \quad (11)$$

where R_{ij} is the geometric distance between the i-th feature point and the j-th feature point in the reference image, and S_{ij} corresponds to the sensed image. Assuming m matches are chosen, $m(m-1)/2$ combinations are calculated totally. The values of the Gong ratio are counted with a statistical method by forming a scale histogram and finding the index of the peak. The feature point pairs contributing to the Gong ratios that are away from the peak index are then found and removed.

In addition, many RANSAC-based methods have been developed for outlier removal. For example, Hossein-nejad and Nasri [9] developed an algorithm so that the shifting threshold is calculated based on the variance. The fast sample consensus (FSC) method proposed by Wu et al. [8] is another approach to improve reliability and efficiency of RANSAC. The FSC algorithm is executed as follows:

1) Randomly choose three correspondences from the matched set.

2) Compute the parameters of the transformation model using these three correspondences.
3) Individually calculate the shifting distances between the transformed moving points and the corresponding fixed points. Collect the feature set if the shifting error is less than one pixel.
4) Repeat Steps 1 to 3. Find the maximal number in the statistical set.

IV. Methods

The objective of the proposed approach is to improve the accuracy of remote sensing image registration by increasing correct matches and removing mismatches. As illustrated in Fig. 1, our image registration method consists of: a) modified SIFT, b) sample feature point matching, c) slope analysis, d) standard grouping boundary, e) confidence elliptical boundary, f) outlier removal, and g) affine transform based on RANSAC.

A. Modified SIFT

The proposed method starts from a modified SIFT using an alternative resampling approach to more accurately find the feature point locations. We down-sampled the basal level image by half using the bilinear interpolation method based on (12) instead of taking every other pixel in each row and column as in the classical SIFT:

$$f(x,y) = f(P_{11})\frac{x_2-x}{x_2-x_1}\frac{y_2-y}{y_2-y_1} + f(P_{21})\frac{x-x_1}{x_2-x_1}\frac{y_2-y}{y_2-y_1} + f(P_{12})\frac{x_2-x}{x_2-x_1}\frac{y-y_1}{y_2-y_1} + f(P_{22})\frac{x-x_1}{x_2-x_1}\frac{y-y_1}{y_2-y_1} \quad (12)$$

where $f(x,y)$ is the intensity value at (x,y) in the spatial domain, and P_{11}, P_{21}, P_{12}, and P_{22} are the locations of the four nearest neighbors.

In addition, after the process of feature point refinement, the coordinates of feature points need to be transformed back to the same scale as the input image. Assuming that the octave index is denoted as zero when the image size in the octave is identical to the input image, we can more accurately obtain the feature point location using

$$f(p,k) = \begin{cases} (p-1)\times 2^k + 2^{k-1} + 0.5 \ , & k > 0 \\ [p \times 2^k] \ , & k \le 0 \end{cases} \quad (13)$$

where p is the found feature point position in one direction (which implies that p must be a positive integer) and k is the index of the octave. Therefore, the transformed coordinates are computed as

$$(x', y') = (f(x,k), f(y,k)) \quad (14)$$

where (x', y') is the transformed coordinate, f is the function in (13), and (x,y) is the coordinate in the original scale.

B. Sample Feature Point Matching

Sample feature point matching adopts the modified SIFT matching process but with a low distance ratio followed by the Gong ratio for statistical counting as shown in Fig. 2. The feature point pairs contributing to the Gong ratios that are away from the peak index are then chosen and removed. The purpose of this procedure is to set the standard grouping boundary (SGB) and the confidence elliptical boundary (CEB) for the second matching.

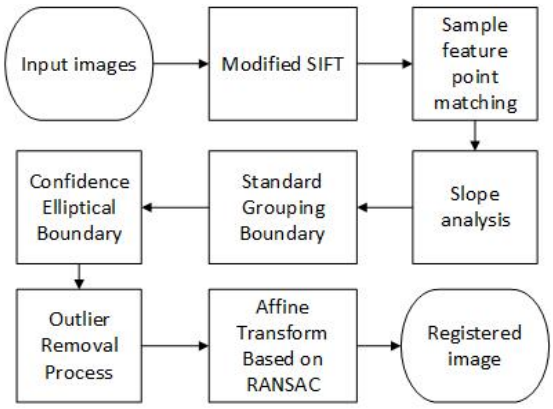


Fig. 1. Main steps of the proposed method.

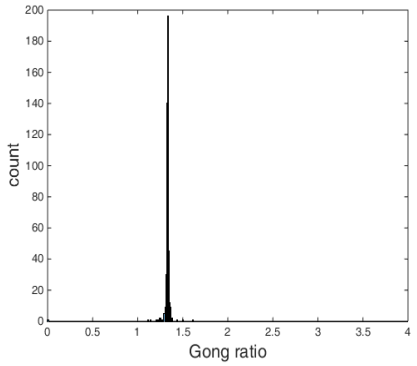


Fig. 2. Illustration of the histogram distribution of the Gong ratio in the sample feature point matching procedure.

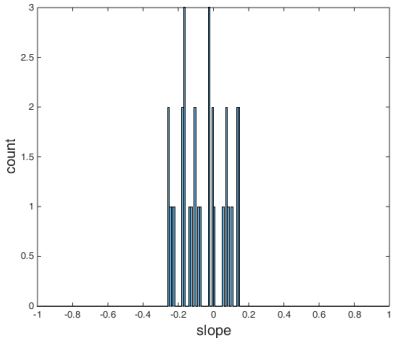


Fig. 3. Distribution of the slope values in the sample feature point matching.

C. Slope Analysis

We then calculate the slopes of the matched pairs using

$$SL_i = \frac{y_R^i - y_S^i}{(x_R^i + \Delta x) - x_S^i} \quad (15)$$

where SL_i is the computed slope, (x_R^i, y_R^i) and (x_S^i, y_S^i) are the i^{th} matched pair coordinates in the reference image and the sensed image, respectively, and Δx is the shifting distance of the reference image, which is set to 512. Besides, the slope SL_i of the matched pairs have several properties:

- Based on the premise of correct matching, the adjacent feature points would have similar values of slopes.
- The slopes in all pairs of feature points would be within a small interval.

The slope histogram distribution of the first matched pairs is shown in Fig. 3, which indicates that all the slope values are within -0.5 to 0.5. Consequently, it is difficult to choose the critical SGB in such a small interval for each group manually. To group those feature points by slopes, the k-means cluster [20] is used to execute data automatically.

Based on these properties, most of the mismatches are removed by setting a boundary to enclose the matches with similar slopes in a specified region. It is possible that a mismatched point located inside the region has a similar slope to other feature points, but its distance is far away from other points. Consequently, the distances of lines have to be taken into consideration by setting a threshold to remove outliers. This distance threshold is defined as 40% of the mean distance using

$$d_{th}^{i} = \frac{\sum_{j=1}^{N} \left\| S_j^i - (R_j^i + X) \right\|}{N} \times 0.4 \tag{16}$$

where d_{th}^{i} is the distance threshold in the i^{th} group, S_j^i and R_j^i are the i^{th} matching pair locations of the sensed image and the reference image, respectively, and X is the shifting distance of the reference image.

D. Standard Grouping Boundary

The concept of the SGB is to set a slope boundary for classifying the feature points in the second matching. The matched pairs are classified into three groups. The SGB is computed based on the following steps.

1) Connect the matched pairs found by the sample feature matching in the sensed image and the shifted reference image with lines.
2) Compute the slopes SL_i in (15) based on the corresponding feature point locations.
3) Classify the feature points in the sensed image into groups based on the slope data using the k-means cluster.
4) Calculate the critical value based on the minimum and maximum slopes in each group.
5) Locate the minimum in Group 1 and the maximum in Group 3.
6) Construct the SGB with a vector containing the four critical values as shown in Fig. 4.

E. Confidence Elliptical Boundary

The purpose of the CEB as shown in Fig. 5 is to remove outliers based on the properties of slopes. Outliers mean wrong matches, which are basically outside the elliptical region. The sample slopes are from the feature point pairs in the first matching. Based on the following steps we can obtain the confidence elliptical function:

$$\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2 = s \tag{17}$$

1) Find the covariance matrix in each group using

$$\Sigma = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} \tag{18}$$

2) Compute the eigenvalues and eigenvectors of the covariance matrix. The largest variance is in the x-axis and the smallest variance is in the y-axis.
3) Compute the angle of the oblique ellipse based on the tangent of the two eigenvectors.

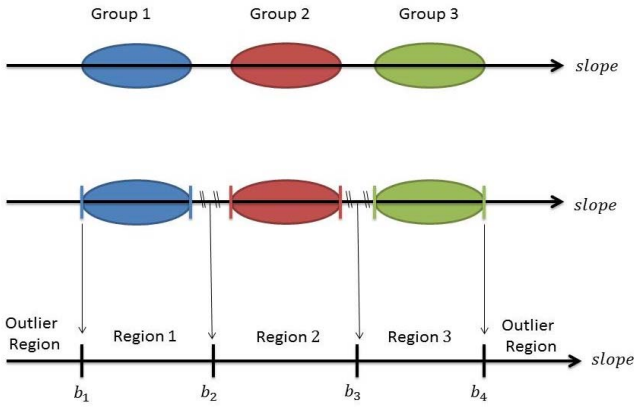


Fig. 4. The sample point slopes are classified by the k-means cluster and the SGB is a boundary vector containing four elements.

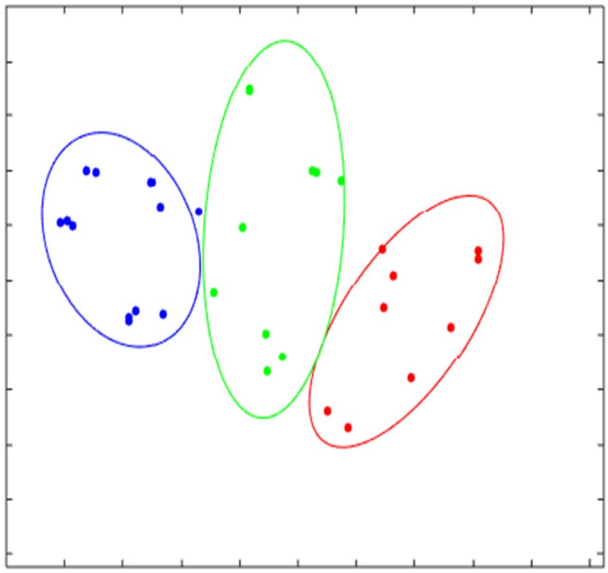


Fig. 5. Illustration of three CEBs in the sensed image.

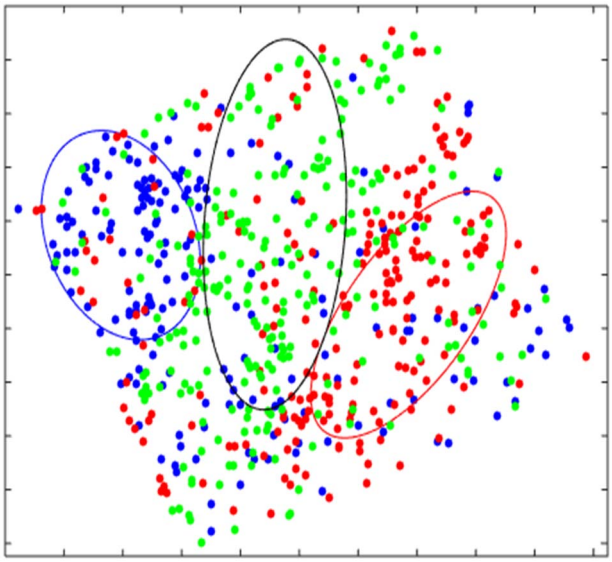


Fig. 6. The matching locations in the sensed image (distance ratio equal to 0.9) are classified by the SGB and the CEB in each group.

4) The center of the ellipse is the mean location of sample feature points and s is 3.24 representing the confidence interval around 85% to 90%.

F. Outlier Removal Process

In this process, the SIFT descriptors of the sensed image and the reference image are matched again (the second matching) with the distance ratio equal to 0.9. To find more correct matches and remove mismatches, both SGB and CEB are utilized as illustrated in Fig. 6. Each feature point pair has its own slope value and all slope values are classified by the SGB. Besides, each group has its own CEB. Based on the feature points in the sensed image, the feature points outside the CEB are removed. To guarantee the adjacent feature points with similar slopes lying inside the CEB, the mean distance of slopes is adopted. The feature point pair whose slope distance is larger than a prescribed threshold is eliminated.

G. Affine Transform Based on RANSAC

In this stage, refined matched pairs are automatically chosen using RANSAC to compute the six parameters of the affine transformation function. To find the six parameters, at least three pairs of matched pairs are required. However, the number of matched pairs is generally much higher than three. The parameters of the affine transformation functions are then obtained by minimizing the sum of the squared distances. Lastly, the registered image is generated based on the affine transform.

V. Experimental Results

A. Design of Experiments

We evaluated the proposed method using the high resolution orthoimagery (HRO) provided by the United State Geological Survey (USGS) [21]. For the design of experiments, we classified remote sensing images into two datasets: the ground truth dataset and multi-temporal image dataset. In the ground truth dataset, sensed images were generated by rotating and translating reference images. The multi-temporal images were taken from the same scene with different timescales. Representative images are illustrated in Fig. 7.

B. Ground Truth Dataset

Numerous image pairs in the ground truth dataset were employed to evaluate the proposed image registration method. The registered results of our framework were compared with the improved SIFT method, which adopted the scheme described in Sec. IV.A only. The MI between the sensed image and the registered image was computed for quantitative comparison. Fig. 8 shows the MI scores of both methods, which indicated that the proposed method is better than the improved SIFT method in that larger MI values were produced.

C. Multi-temporal Image Dataset

In practical applications of remote sensing images, the use of registration is to align two multi-temporal images and demonstrate the difference between these two images. An example is illustrated in Fig. 9. Fig. 9(a) depicts the feature

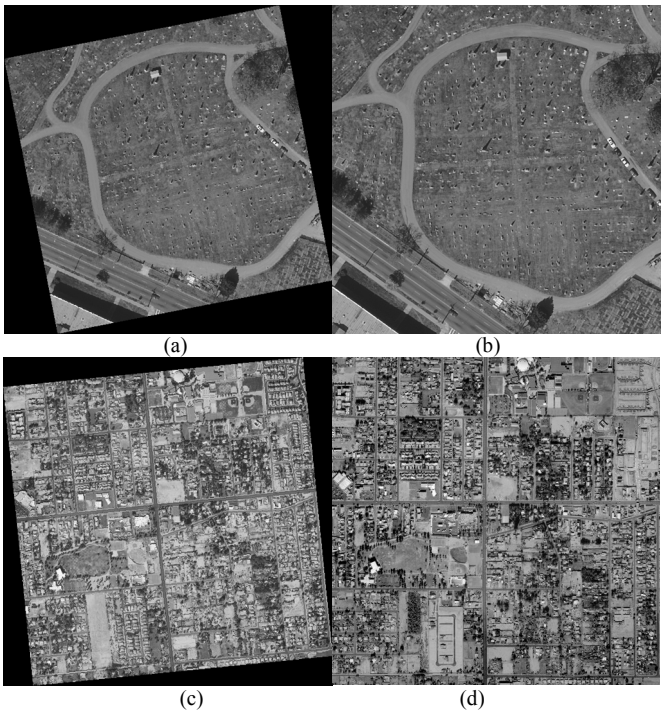


Fig. 7. Representative image pairs. (a) and (b) are the pair of test images. (c) and (d) are the pair of multi-temporal images.

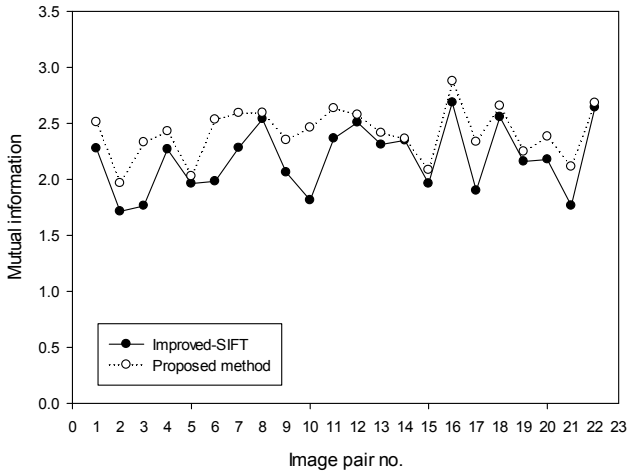


Fig. 8. Comparison in registration accuracy between the improved SIFT method and the proposed method.

point pair results of using the distance ratio equal to 0.9. Based on the CEB scheme, correct matches were effectively extracted as shown in Fig. 9(b). Subsequently, as depicted in Fig. 9(c), the RANSAC method further selected the most suitable matches for the transformation function. Table I summarizes the quantitative evaluation of different methods based on the RMSE index. It is obvious that the proposed framework achieved the lowest RMSE values in all illustrations, which indicated better registration accuracy. This demonstrated the advantages of the modified SIFT procedure in detecting feature points more accurately. The registered image was compared with the reference image for better understanding as shown in Fig. 10. The red regions (new buildings) in Fig. 10(b) indicated the difference between the two images.

TABLE I
COMPARISON IN RMSE BETWEEN DIFFERENT METHODS

Image	Method	a_1	a_2	a_3	a_4	δ_1	δ_2	RMSE	Parameter
irport	SIFT	0.7119	-0.2857	0.2850	0.7090	147.0794	0.707	0.0812	dR=0.5
	SIFT+Gong ratio	0.7105	-0.2851	0.2838	0.7096	146.763	1.3589	0.0777	dR=0.7, th=50
	SIFT+FSC	0.7108	-0.2851	0.2840	0.7094	146.8751	1.3039	0.0766	dR=0.8
	Proposed	**0.7086**	**-0.2854**	**0.2840**	**0.7108**	**147.739**	**1.3706**	**0.0544**	**dR1=0.4**
Bar	SIFT	0.5811	-0.4424	0.5785	0.6083	213.8370	-40.3101	2.6632	dR=0.4
	SIFT+Gong ratio	0.52	-0.4919	0.5732	0.6072	226.6482	-24.3933	0.3163	dR=0.6,th=10
	SIFT+FSC	0.5081	-0.4880	0.5789	0.6047	225.7483	-21.7401	0.3447	dR=0.5
	Proposed	**0.6095**	**-0.4954**	**0.5742**	**0.5168**	**228.0773**	**-23.7435**	**0.1653**	**dR1=0.4**
Cape	SIFT	0.6259	-0.5163	0.4855	1.0033	125.8775	-29.907	4.5857	dR=0.6
	SIFT+Gong ratio	0.6673	-0.3438	0.4279	0.7458	152.5805	-24.686	0.2912	dR=0.7,th=10
	SIFT+FSC	0.51	-0.4876	0.5794	0.6048	225.7428	-22.366	0.3107	dR=0.6
	Proposed	**0.7453**	**-0.3419**	**0.4258**	**0.6689**	**152.9782**	**-24.5167**	**0.1328**	**dR1=0.4**
Crossing	SIFT	0.5083	-0.4772	0.4815	0.5142	246.2726	1.1425	0.1726	dR=0.6
	SIFT+Gong ratio	0.5086	-0.4758	0.4805	0.5141	245.9256	1.1176	0.1697	dR=0.8,th=10
	SIFT+FSC	0.5084	-0.4750	0.4806	0.5131	245.9404	1.0414	0.1799	dR=0.7
	Proposed	**0.5145**	**-0.4747**	**0.4816**	**0.5091**	**246.5216**	**0.8349**	**0.0893**	**dR1=0.4**

dR is the distance ratio in the SIFT matching process. dR1 is the distance ratio in the sample feature point matching process. th is the threshold in the Gong process.

VI. CONCLUSION

One objective of remote sensing image registration is to display the difference between multi-temporal images for further analysis. We have developed an effective algorithm for automatic remote sensing image registration to achieve this goal. One advantage of the proposed method is that only one parameter of the distance ratio is required for image registration. This parameter is robust in that all experimental images were executed using a constant value and no iterative process is needed. The proposed algorithm combined a series of techniques including the modified SIFT, the Gong ratio, the elliptical boundaries, and the RANSAC methods. Experimental results demonstrated the advantages of our image registration framework over stat-of-the-art methods. We believe that this new scheme is of potential in many remote sensing applications that require automatic image registration.

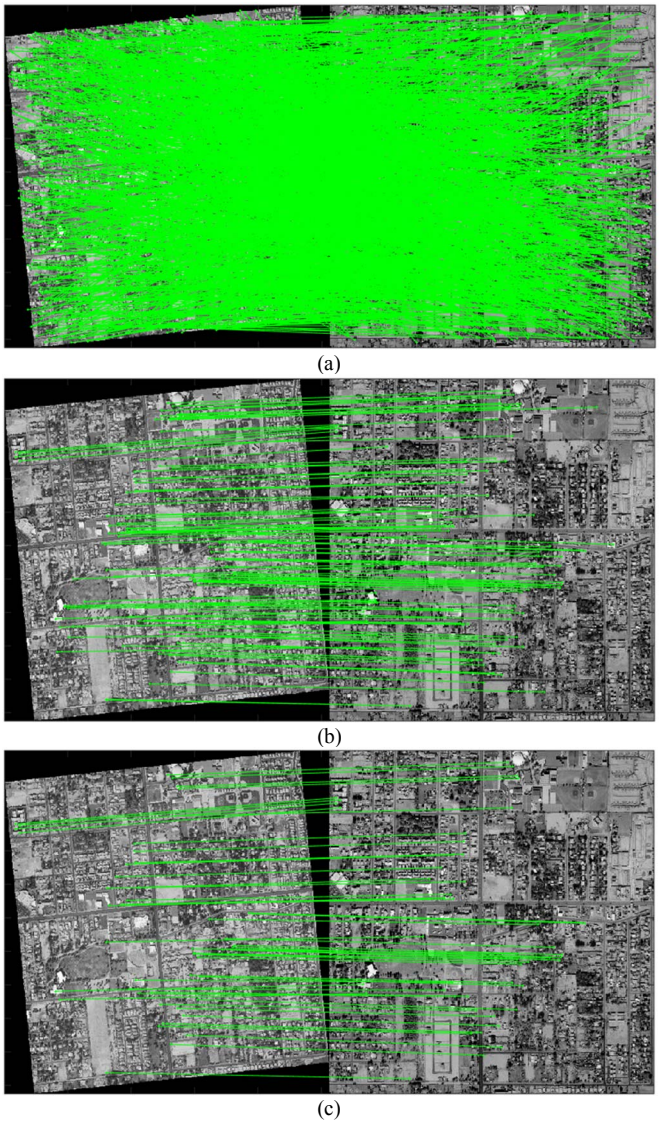


Fig. 9. (a) Matched results with the distance ratio equal to 0.9. (b) Matched pairs after the outlier removal process with the CEB. (c) Final matched pairs after the RANSAC procedure for the transformation function.

REFERENCES

[1] H. E. M. Meier, A. Hoglund, K. Eilola, and E. Almroth-Rosell, "Impact of accelerated future global mean sea level rise on hypoxia in the Baltic Sea," (in English), *Climate Dynamics,* vol. 49, no. 1-2, pp. 163-172, Jul 2017.

[2] J. L. Moigne, N. S. Netanyahu, and R. D. Eastman, *Introduction* (Image Registration for Remote Sensing). UK:Cambridge University Press, 2011.

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision,* vol. 60, no. 2, pp. 91-110, 2004.

[4] Q. Li, G. Wang, J. Liu, and S. Chen, "Robust scale-invariant feature matching for remote sensing image registration," *IEEE Geoscience and Remote Sensing Letters,* vol. 6, no. 2, pp. 287-291, 2009.

[5] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 52, no. 7, pp. 4328-4338, 2014.

[6] B. Li and H. Ye, "RSCJ: Robust sample consensus judging algorithm for remote sensing image registration," *IEEE Geoscience and Remote Sensing Letters,* vol. 9, no. 4, pp. 574-578, 2012.

[7] C. Huo, C. Pan, L. Huo, and Z. Zhou, "Multilevel SIFT matching for large-size VHR image registration," *IEEE Geoscience and Remote Sensing Letters,* vol. 9, no. 2, pp. 171-175, 2012.

[8] Y. Wu, W. Ma, M. Gong, L. Su, and L. Jiao, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geoscience and Remote Sensing Letters,* vol. 12, no. 1, pp. 43-47, 2015.

[9] Z. Hossein-nejad and M. Nasri, "Image registration based on SIFT features and adaptive RANSAC transform," in *Communication and Signal Processing (ICCSP), 2016 International Conference on*, 2016, pp. 1087-1091: IEEE.

[10] W. Ma *et al.*, "Remote Sensing Image Registration With Modified SIFT and Enhanced Feature Matching," *IEEE Geoscience and Remote Sensing Letters,* vol. 14, no. 1, pp. 3-7, 2017.

[11] P. R. Beaudet, "Rotationally invariant image operators," in *Proc. 4th Int. Joint Conf. Pattern Recog, Tokyo, Japan, 1978*, 1978.

[12] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey vision conference*, 1988, vol. 15, no. 50, p. 10.5244: Manchester, UK.

[13] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London B: Biological Sciences,* vol. 207, no. 1167, pp. 187-217, 1980.

[14] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, 1999, vol. 2, pp. 1150-1157: Ieee.

[15] D. Russakoff, C. Tomasi, T. Rohlfing, and C. Maurer, "Image similarity using mutual information of regions," *Computer Vision-ECCV 2004,* pp. 596-607, 2004.

[16] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE transactions on medical imaging,* vol. 22, no. 8, pp. 986-1004, 2003.

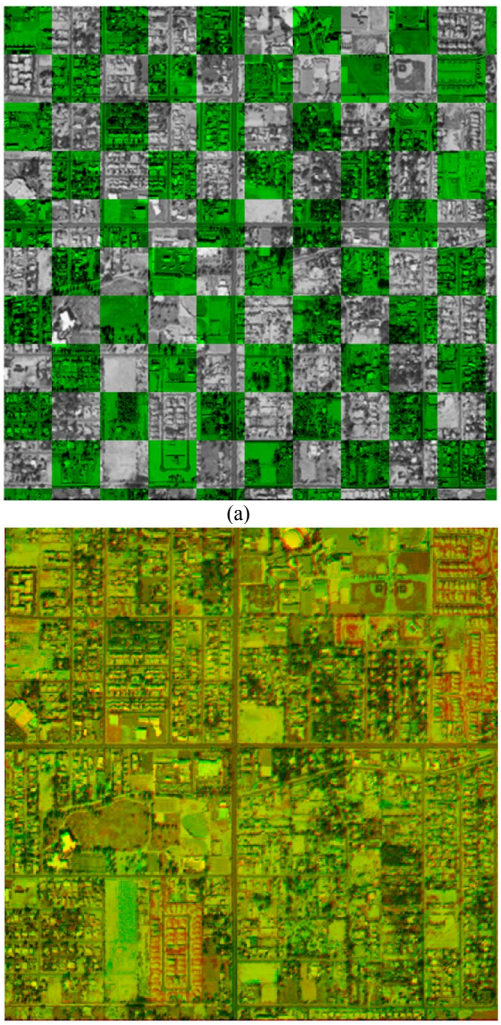
[17] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review,* vol. 5, no. 1, pp. 3-55, 2001.

[18] M. Brown and D. G. Lowe, "Invariant Features from Interest Point Groups," in *BMVC*, 2002, vol. 4.

[19] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM,* vol. 24, no. 6, pp. 381-395, 1981.

[20] T. Garg and A. Malik, "Survey on various enhanced K-means algorithms," *International Journal of Advanced Research in Computer and Communication EngineeringVol,* vol. 3, 2014.

[21] USGS.gov. Earth Explorer [Online]. Available: https://earthexplorer.usgs.gov/



(a)

(b)

Fig. 10. (a) Checkerboard mosaicked images using the proposed method. (b) Comparison between the sensed image and the registered image in the color mode.