

# A Fast and Robust Scene Matching Method for Navigation

Sanhai Ren<sup>1, \*</sup> and Wenge Chang<sup>2</sup>

**Abstract**—The selection of matching method is critical to the scene matching navigation system, as it determines the accuracy of navigation. A coarse-to-fine matching method, which combines the area-based and feature-based matching method, is presented to meet the requirements of navigation, including the real-time performance, the sub-pixel accuracy and the robustness. In the coarse matching stage, the real-time performance is achieved by a pyramid multi-resolution technique, and the robustness is improved by multi-scale circular template fusion. In the precise matching stage, an improved SIFT method is introduced to calculate the matching position and the rotation angle. To validate the method, some experiments are completed. The results show that the proposed method can achieve the sub-pixel matching accuracy and improve the angle accuracy to  $0.1^\circ$ .

## 1. INTRODUCTION

The INS/SAR navigation system, which makes use of synthetic aperture radar (SAR) technology, not only has all-weather and all-day capabilities, but also increases the capabilities of interference suppression. The navigation system matches a real-time image and a geo-referenced image (i.e., reference image) to obtain the position update information, and corrects the cumulated error of Inertial Navigation System (INS) by fusing the updated information with current INS estimates [1]. As one of the key components of the system, scene matching method must satisfy three requirements: the real-time performance, the sub-pixel accuracy and the robustness.

Based on the gray-scale of each pixel, conventional matching methods focus on the correlation between reference image and real-time image. The major steps of these methods are sliding one image over another, and calculating the degree of similarity at each pixel. The computation time of these methods is too long to navigation, since the methods demand searching over a whole image. Moreover, they are sensitive to image rotation and scaling change, and can only achieve pixel-level accuracy.

Tremendous research attentions have been attracted to overcome the drawbacks of traditional methods. Some of these methods are aiming at speeding up the matching process [2–4], while some others are aiming at improving the robustness [5, 6]. Most existing drawbacks have been extensively solved subjected to given backgrounds, however, these methods have not reached a point that can be easily adopted into a practical system and expected to perform reliably well on various backgrounds.

More recently, the methods underpinned by local descriptors, i.e., Scale Invariant Feature Transform (SIFT), have been employed in image matching widely [7–9]. This approach is regarded as one of the robust feature-based matching methods against the rotation, scaling, and illumination change. Mikolajczyk and Schmid [10] evaluated the performances of various local descriptors, and found that SIFT outperforms other descriptors for feature matching. However, the computation time will sharply increase when it is applied in images of large size. Moreover, this method is proposed for optical image matching, and issues may arise when it is used in SAR image matching. For instance, the SIFT keypoints extracted from SAR image pairs might be too sparse to calculate transformation parameters and the ratio of false matches will increase greatly.

---

*Received 13 March 2014, Accepted 18 April 2014, Scheduled 15 May 2014*

\* Corresponding author: Sanhai Ren (sanhairen@163.com).

<sup>1</sup> Beijing Institute of Tracking and Telecommunication Technology, Beijing 100094, China. <sup>2</sup> National University of Defense Technology, Hunan, Changsha 410073, P. R. China.

In the SAR scene matching navigation system, enough features must be covered by a real-time image to ensure a correct matching, which results in a large size of both the real-time image and the reference image. Furthermore, the real-time image and the reference image may be acquired from different incidence angles, at different times and with different resolutions. Therefore, it is impossible for traditional methods to fulfill the real-time performance and robustness simultaneously. In this article, we present a fast and robust matching method for navigation. It combines the techniques of pyramid multi-resolution decomposition and multi-scale circular template fusion to achieve the real-time performance and the robustness. Following that, a set of accurate tie-points are detected by an improved SIFT method and used to calculate the position error and the yawing error, which achieves the sub-pixel accuracy.

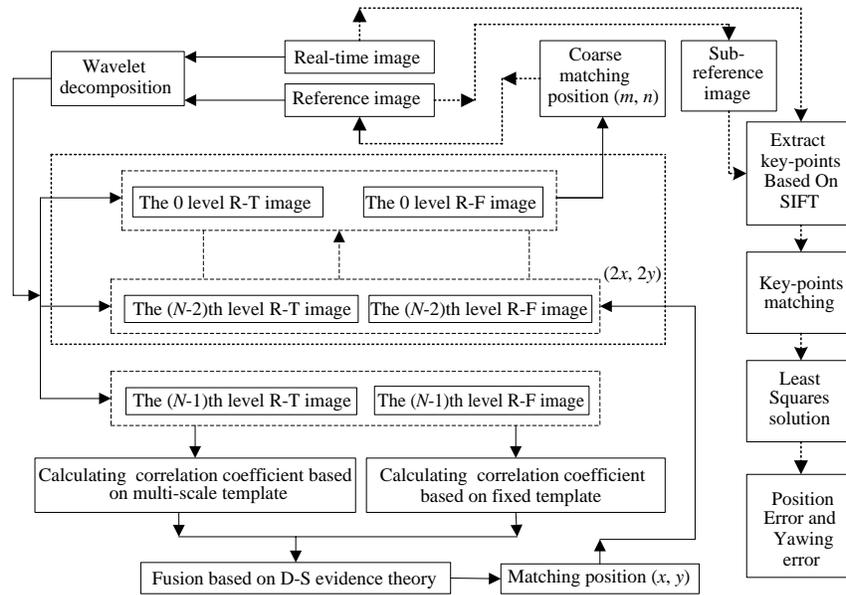
This article is organized as follows. An overview of the proposed matching method is presented in Section 2. The coarse matching method is described in Section 3, which mainly deals with the fast matching scheme, the matching template selection, and the construction of multi-scale template. The problems and solutions when SIFT matching method is used in SAR image are discussed in Section 4. In order to validate the method, plentiful experiments are completed in Section 5. Finally, the related conclusions are given in Section 6.

## 2. THE OVERVIEW

An overview of the proposed method is demonstrated in Figure 1, where the arrowhead of solid line represents the flow of coarse matching, and the arrowhead of dashed line represents the flow of precise matching.

At the start of the method, the real-time image and reference image are both decomposed into  $N - 1$  levels. At each level, the matching position is acquired via the technique of multi-scale circular template fusion. The matching position in the original image is called the coarse matching position, based on which a sub-reference image with the same size of the real-time image can be intercepted from the reference image.

Then, SIFT features in the real-time image and sub-reference image are extracted via an improved SIFT method. The best candidate match for each key-point in the real-time image is found by identifying its nearest neighbor in the database of key-points from the sub-reference image, where the nearest neighbor is defined as the key-point with minimum Euclidean distance for the descriptor vectors [8].



**Figure 1.** The flow chart of the proposed method, where R-T image is the real-time image, R-F image is the reference image.

In the end, a set of accurate tie-points are acquired from an improved SIFT method, and used for calculating the position error and the yawing error.

### 3. THE COARSE MATCHING METHOD

#### 3.1. The Fast Matching Scheme

In order to speed up the matching method, a wavelet-based pyramid multi-resolution technique and a coarse-to-fine search strategy is employed, which improves the search efficiency greatly.

The constructed pyramid is demonstrated in Figure 2.  $L_0-A$  and  $L_0-B$  are the primal reference image and real-time image, while  $L_1-A$ ,  $L_1-B$  and  $L_2-A$ ,  $L_2-B$  are the corresponding 1-level and 2-level wavelet transform. After wavelet transform, a high resolution image can be decomposed into four parts: one low frequency sub-band (LL) and three high frequency sub-bands (LH, HL, HH). The low frequency part will be decomposed further, and iteratively as required. The translation sensitivity of wavelet-based image matching was investigated in [11]. The results show that the low frequency sub-band image of wavelet decomposition is better than the high frequency sub-band in the insensitivity to translation. Hence, a better matching effect based on low frequency sub-band can be obtained.

The match starts from the LL sub-band (i.e., LL2 in the  $L_2-A$  and the  $L_2-B$  of Figure 2) of the top level Discrete Wavelet Transform (DWT), and then iteratively for sub-bands on each level going down, till the best match position in the corresponding search spaces is found. At the top level of

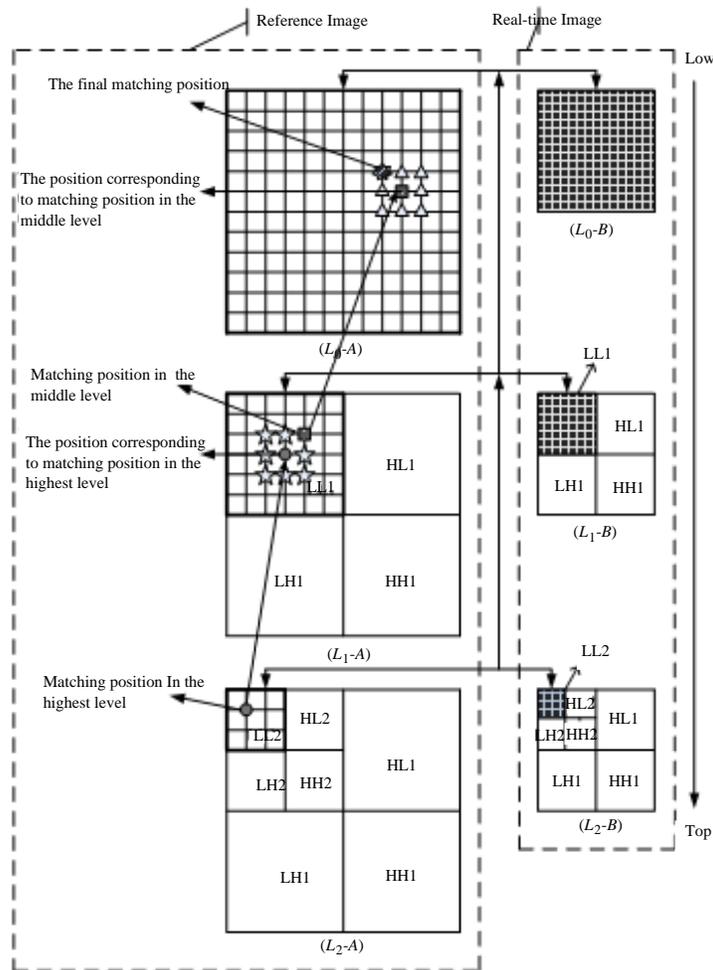


Figure 2. The matching method based on pyramid multi-resolution.

DWT, the search is performed exhaustively over the whole search space. At the next level of DWT, only  $[2T_m - R, 2T_m + R]$  is selected as the search space, where  $T_m$  is the search result of translation offset in the previous step, and  $R$  is a positive constant representing the search range.

### 3.2. The Robust Matching Scheme

In the practical application, we find that the matching method based on gray-scale and correlation is not only sensitive to image rotation, but also sensitive to scaling change. Therefore, some methods based on invariant moments are introduced to overcome it [12]. However, these moment-based methods lose their functions here, since the wavelet transform used in our method will change the distribution of gray-level.

Some matching methods decrease the sensitivity to scaling change by extracting features in the scale space, and achieve satisfactory results [7–9]. Therefore, a new method is presented to decrease the sensitivity to rotation and scaling change in this article, using the fusion of multi-scale circular template. For the fusion strategy, the Dempster-Shafer (D-S) evidence theory is employed [13].

#### 3.2.1. The Fusion Strategy Based on D-S Evidence Theory

D-S evidence theory, a statistical data fusion classification method, has the advantage of presenting and dealing with uncertain information. Knowledge from multiple information sources about events (called propositions) is combined by Dempster's rule to reduce the interval of uncertainty, and find the intersection or conjunction of the propositions, as well as the associated probability.

The detailed definitions of the basic concepts of D-S evidence, such as the frame of discernment, Basic Probability Assignment (BPA), Basic Probability Number (BPN), focal element, Belief Function (Bel) and Plausibility Function (Pl) can be found in [14]. Suppose  $\text{Bel}_1$  and  $\text{Bel}_2$  are two Bels of the same frame of discernment  $\mathbf{X}$ ,  $m_1$  and  $m_2$  are the corresponding probability assignment function,  $A_1, A_2, \dots, A_k$  and  $B_1, B_2, \dots, B_l$  are focal elements. Then the BPA ( $m(A)$ ) of focal element  $A$  can be calculated by (1).

$$m(A) = \begin{cases} \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)}, & \forall A \subseteq X, A \neq \emptyset \\ 0, & A = \emptyset \end{cases} \quad (1)$$

where  $\emptyset$  is empty set.

In the scene matching, there are only two elements in the frame of discernment: the matching element ( $\mathbf{M}$ ) and the mismatching element ( $\mathbf{N}$ ). Therefore, the frame of discernment is  $\mathbf{X} = \{\mathbf{M}, \mathbf{N}\}$ . For each element in the output sequence, if the Bel based on proposition 1 is  $m_1(\mathbf{M})$ , then its Pl can be calculated by  $m_1(\mathbf{X}) = 1 - m_1(\mathbf{M})$ . Similarly, the  $m_2(\mathbf{M})$  and  $m_2(\mathbf{X})$  can be got. Thus, the associated probability based on both of the proposition 1 and the proposition 2 is

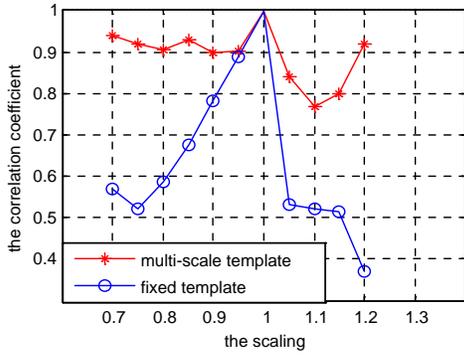
$$m(\mathbf{M}) = m_1(\mathbf{M}) m_2(\mathbf{M}) + m_1(\mathbf{M}) m_2(\mathbf{X}) + m_1(\mathbf{X}) m_2(\mathbf{M}) \quad (2)$$

According to the D-S evidence theory, the Bel of each element should be known in order to fulfill the fusion. The correlation coefficient, which denotes the degree of similarity between the real-time image and the corresponding sub-reference image, can be considered as the Bel. If the correlation coefficient is  $\rho_1(i, j)$  based on one template, and  $\rho_2(i, j)$  based on another template, then the associated correlation coefficient  $\rho(i, j)$  can be calculated as

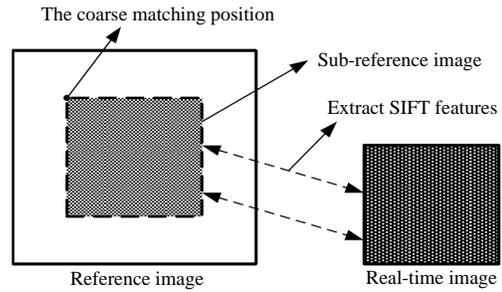
$$\rho(i, j) = 1 - (1 - \rho_1(i, j))(1 - \rho_2(i, j)) \quad (3)$$

#### 3.2.2. Construction of Multi-Scale Circular Template

The key of the method is to construct the multi-scale circular templates. Generally, the matching accuracy will be higher if the scaling ratio of multi-scale template is more similar to that of the image pair. However, the ratio is always unknown in the practical application. If too many multi-scale templates are constructed to approach the real value, the storage and computation loads will rise



**Figure 3.** The relationship between scaling and correlation coefficient.



**Figure 4.** The real-time image and sub-reference image.

sharply. Luckily, the resolutions of image pairs are known, hence the ratio of resolutions is taken as the approximate ratio of scaling in this article.

Suppose the size of the primal real-time image is  $M_r \times N_r$ , and the scaling ratio is  $s_{ca}$ . If  $s_{ca} < 1$ , the real-time image at each level of the pyramid is up-sampled, which results in an enlarged image with size of  $fix(M_r/s_{ca}) \times fix(N_r/s_{ca})$ , where  $fix(\cdot)$  represents rounding the elements toward zero. Then, a new real-time image with size of  $M_r \times N_r$  is constructed by selecting the medial parts of the enlarged image, and the corresponding circular template can be constructed. In this situation, the size of the constructed template is the same as the fixed template.

Otherwise, if  $s_{ca} > 1$ , the real-time image at each level of the pyramid is down-sampled, and a zoomed-out image with size of  $fix(M_r/s_{ca}) \times fix(N_r/s_{ca})$  is acquired. For each sub-reference image, a new sub-reference image with the size of  $fix(M_r/s_{ca}) \times fix(N_r/s_{ca})$  is constructed by selecting its medial parts, and the corresponding circular template can be constructed. Thus, the size of the constructed template is smaller than that of the fixed template.

In the end, the correlation coefficient of the fixed template and the multi-scale template are calculated respectively, and then combined based on (3).

An experiment is completed to demonstrate the insensitivity to scaling change of multi-scale circular template, as in Section 3.2. The results are shown in Figure 3, from which we can get that the matching method can decrease the sensitivity to scaling change after the fusion of multi-scale circular template. In other words, the method increases its adaptability to geometric distortions.

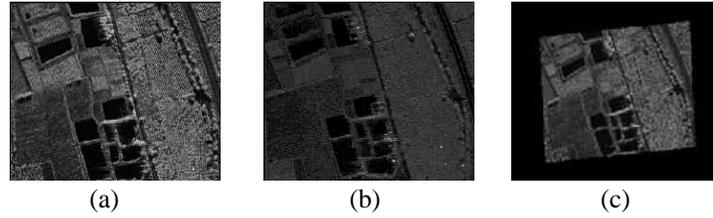
#### 4. THE PRECISE MATCHING METHOD

Based on the coarse matching position, we can get a sub-reference image which has the same size as the real-time image, as illustrated in Figure 4. In order to estimate the rotation angle and matching position accurately, a set of tie-points between real-time image and sub-reference image are needed. In this article, a kind of SIFT-based detector is selected as the point detector. The speediness methods of SIFT, such as integral image [14] and Principal Component Analysis (PCA) methods [15] are not taken into account, since the size of sub-reference image is much smaller than that of the original reference image. The stress is just laid on the problems and solutions when SIFT is used in SAR images.

##### 4.1. Limitations of SIFT Used in SAR Images

SIFT detects local extrema in DOG scale space, making it robust against additive noise. However, noise in SAR image follows multiplicative noise model, and cannot be mitigated by DOG processing. For instance, for the given image pair in Figure 5(a) and Figure 5(b), the original SIFT achieves only 4 matches, which are not sufficient for parameter calculation, because at least 3–4 times of the least matches (2 matches) are needed for reliability.

SIFT assigns a 128-element feature vector to each key-point as its descriptor. For each key-point in the real-time image, the Euclidean distance between it and all key-points in the reference image is



**Figure 5.** Example image pairs. (a) The SAR image of Ku band. (b) The SAR image of L band. (c) The transformed image from (a), where the scale factor is 0.7 and rotation angle is  $4^\circ$ .

calculated firstly, and then the tie-point is selected by judging whether the Euclidean distance of the closest neighbor to that of the second-closest neighbor is less than a threshold [8]. This is very effective for optical image, but not for SAR image. For example, there are 40 false matches out of 110 matches in Figure 5(b) and Figure 5(c), where 12 of these false matches have a scale ratio (defined in Formula (5)) less than 0.59 or greater than 0.805, and 35 of these false matches have a rotation difference (defined in Formula (6)) less than  $2^\circ$  or larger than  $6^\circ$ , which are beyond the ranges for scale ratio and rotation differences as defined in the step 4 of Section 4.3.

Therefore, when SIFT is used in SAR image, it might be not able to detect enough tie-points, or the ratio of false matches to total matches is too high.

#### 4.2. Previously Proposed Solutions

The methods to eliminate false matches are exploited in two independent studies [16, 17]. Both of these methods use an eliminate-after-matching approach, hence they intend to reduce false matches in the SIFT output. Yi et al. [16] form a histogram of scale difference and define a window around the peak of this histogram. The matches with scale difference outside this window are rejected. A limitation in the method is that only images pairs with approximately the same scale are considered. Bastanlar et al. [17] make an improvement on the method proposed in [16]. They preprocess image pairs using scale ratio to adjust their scales, and then eliminate false matches as [16] does. Compared with the method proposed in [16], the method in [17] can not only eliminate false matches, but also increase the number of correct matches. However, the computation time is too long to be adopted in the navigation system. Moreover, both of them are proposed for optical image, and the performance will decrease when used in SAR image.

#### 4.3. The Proposed Solution

According to the characters of SIFT, the SAR image pairs are preprocessed to transform the noise model, and matched by SIFT. Then, the false matches are eliminated by using the dominate scale ratio and rotation difference, from which a significant improvement on SIFT matching can be achieved.

If we use  $Y$  to denote the SAR image amplitude, and  $X$  is the speckle-free value of the magnitude of the electric field of the backscattered signal, then  $Y$  is related to  $X$  by the multiplicative model  $Y = F \cdot X$ . For a single-look image, the normalized fading random variable  $F$  obeys a Rayleigh distribution [18].

In optical image,  $Y$  is related to  $X$  by an additive model  $Y = F + X$ , and  $F$  follows the Gaussian distribution. In order to make SIFT more suitable for SAR images, logarithmic transformation is employed to convert the multiplicative noise model to an additive noise model.

After a natural logarithmic transformation, SAR images can be converted into

$$\ln(Y) = \ln(F) + \ln(X) = \hat{F} + \hat{X} \quad (4)$$

where  $\hat{F}$  follows the double exponential or Fisher-Tipper distribution.

It has been established that as the number of SAR looks increase, the speckle random variable approaches a Gaussian distribution. In [18], a distance between cumulative distributions is computed to measure the deviation of the log-transformed speckle from Gaussianity. It is clearly demonstrated

that for amplitude image, the log-transformed speckle noise is already statistically very close to the Gaussian distribution. A conclusion was drawn that the Gaussian approximation is relatively good when  $L \geq 3$ .

Consequently, the noise in the log-transformed SAR image can be considered as additive noise, and its Probability Density Function (PDF) will be more similar to Gaussian distribution with the increasing number of SAR looks. One issue related to the logarithmic transformation arises from the fact that the mean of the log-transformed speckle noise is not zero, whereas a significant set of techniques assume Gaussian white noise with zero mean. Therefore, we should subtract the mean value from the log-transformed image to make the noise with zero mean.

Each key-point detected by SIFT is described by a vector  $(x_i, y_i, \sigma_i, \theta_i)$ . Elements of the vector denote location  $(x_i, y_i)$ , scale  $\sigma_i$  and orientation  $\theta_i$ . The definition of scale ratio ( $S_R$ ) and rotation difference ( $R_D$ ) of key-point pair  $P_1(x_{1i}, y_{1i}, \sigma_{1i}, \theta_{1i})$  and  $P_2(x_{2i}, y_{2i}, \sigma_{2i}, \theta_{2i})$  are:

$$S_{Ri}(P_1, P_2) = \sigma_{1i}/\sigma_{2i} \quad i = 1, 2, \dots, k \quad (5)$$

$$R_{Di}(P_1, P_2) = \theta_{1i} - \theta_{2i} \quad i = 1, 2, \dots, k \quad (6)$$

where  $k$  is the number of correct matches.

For correct matches,  $S_R$  and  $R_D$  should be similar or equal to the real values, as the SIFT scale space ratio and rotation difference reveal the scale ratio and rotation difference of the features between the images. Therefore, the matches with large  $S_R$  and  $R_D$  deviation from the real value should be eliminated.

In order to get the approximate scale ratio and rotation difference, histograms of  $S_R$  and  $R_D$  are formed respectively. Then the peak and two values closest to it of each histogram are gotten. Finally, a parabola is fit to the three histogram values to interpolate the peak position for better accuracy.

The improved SIFT-based matching method contains five steps, summarized as follows:

Step 1: Preprocess the image pairs using a logarithmic transformation to convert the noise model from a multiplicative model to an additive model.

Step 2: Perform SIFT matching on the preprocessed image pairs, and plot histograms of  $S_R$  and  $R_D$  of tie-points.

Step 3: Extract the peaks of each histogram, and take them as the approximate scale ratio  $P_s$  and rotation difference  $P_r$  of the images to be matched. Then, take the  $P_s$  and  $P_r$  as the real value of the scale ratio and rotation difference of the matches.

Step 4: Eliminate the false matches whose  $S_R$  are beyond the scope  $[0.8 \times P_s, 1.2 \times P_s]$ , or  $R_D$  are beyond the scope  $[P_r - 2^\circ, P_r + 2^\circ]$ , where the scopes are acquired from statistical experiments [16].

Step 5: Calculate the translation and rotation angle between the real-time image and the reference image through the least-squares solution.

## 5. EXPERIMENTAL ANALYSIS

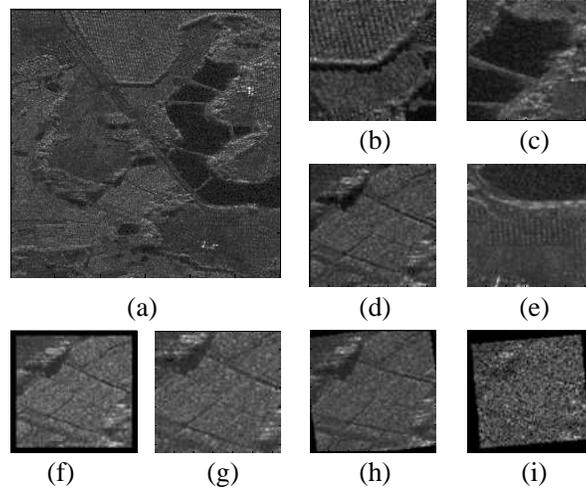
### 5.1. Create Image Pairs and Solve Transformation Parameters

A large number of image pairs are required to perform simulation experiments on computer. If image pairs already exist, serial segments of real-time images could be produced directly from them. However, sometimes there are not enough image pairs available but only a single image. Thus segments of real-time images can be created by cutting casually from the reference image with some artificial distortions. An example of such simulation is shown in Figure 6.

An affine transformation is adopted to form distortions of the sub-images intercepted from reference image. The affine transformation from a reference image point  $[x_i, y_i]$  to a real-time image point  $[x_o, y_o]$  can be expressed as:

$$\begin{bmatrix} x_o \\ y_o \end{bmatrix} = S \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (7)$$

where the model translation is  $[t_x, t_y]$  and the affine rotation and scaling are represented by  $\theta$  and  $S$ . Then, the values of pixels can be calculated by bilinear interpolation.



**Figure 6.** Example of created image pairs. (a) The reference image. (b)~(e) The  $128 \times 128$  pixel images cut from (a), from which the real-time image will be simulated with some distortions. (f)~(i) The simulation segments of real-time image from (d), where (f) is the image zoomed out by 10%, (g) is the image zoomed in by 20%, (h) is the image rotated by  $8^\circ$ , (i) is the image zoom out by 10%, rotated by  $5^\circ$ , and added noise with variance 2.

Suppose there are  $k$  tie-points, then (7) can be rewritten to gather the unknown variables into a column vector:

$$\begin{bmatrix} x_{o1} \\ y_{o1} \\ \vdots \\ x_{ok} \\ y_{ok} \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_{i1} & -y_{i1} \\ 0 & 1 & y_{i1} & x_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{ik} & -y_{ik} \\ 0 & 1 & y_{ik} & x_{ik} \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ s \cos \theta \\ s \sin \theta \end{bmatrix} \quad (8)$$

where  $k$  is the number of correct matches. At least 2 matches are needed to provide a solution. Generally, in order to improve the accuracy, at least 3–4 times of the least matches (6–8 matches) are needed.

$$\text{Let } \mathbf{A} = \begin{bmatrix} x_{o1} \\ y_{o1} \\ \vdots \\ x_{ok} \\ y_{ok} \end{bmatrix}, \mathbf{H} = \begin{bmatrix} 1 & 0 & x_{i1} & -y_{i1} \\ 0 & 1 & y_{i1} & x_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{ik} & -y_{ik} \\ 0 & 1 & y_{ik} & x_{ik} \end{bmatrix}, \mathbf{P} = \begin{bmatrix} t_x \\ t_y \\ s \cos \theta \\ s \sin \theta \end{bmatrix}, \text{ then we can rewrite (8) as}$$

$$\mathbf{A} = \mathbf{H} \cdot \mathbf{P} \quad (9)$$

Then, the unknown variables can be determined by the least-squares solution.

$$\mathbf{P} = [\mathbf{H}^T \mathbf{H}]^{-1} \mathbf{H}^T \mathbf{A} \quad (10)$$

The translation  $[t_x, t_y]$  can be obtained directly, and the rotation angle can be calculated by solving the following equation.

$$\theta = \tan^{-1} (\mathbf{P}(4) / \mathbf{P}(3)) \quad (11)$$

## 5.2. Experimental Analysis

In the statistical experiments, the results are shown by experimental indices. The matching probability, matching error and matching time are taken as capability indices in this article [19].

- 1) Matching probability ( $P_a$ ): the ratio of correct matching times ( $N_c$ ) and total matching times ( $N_a$ )

$$P_a = N_c / N_a \quad (12)$$

2) Matching error ( $\sigma$ ): the mean average of distance between calculated value and theoretic value.

$$\sigma = \sum_{i=1}^{N_c} \varepsilon_i / N_c \tag{13}$$

3) Matching time ( $T$ ): the median time of a matching algorithm completing once account.

$$T = \sum_{i=1}^{N_c} t_i / N_c \tag{14}$$

Figure 6(a) ( $512 \times 512$  pixels) is a selected reference image. In order to simulate real-time images, blocks of sub-images with  $128 \times 128$  pixels are intercepted casually from it, as shown in Figures 6(b)–(e). Then, distortions and speckle noise are added. Figures 6(f)–(i) show some examples of simulative real-time images. The statistical results of the method and other matching methods are shown in Table 1, where the matching with error less than 3 pixels is considered as a correct match. The angle accuracy is the mean angle error between the calculated value and the real one.

All the methods are implemented in MATLAB on a PC with Intel Pentium 4/2.4 GHz processor and 1 GB memory. From the data in Table 1, we can get that the performance of the proposed method is perfect both in single distortion and combined distortion (zoom in by 10%, rotated by  $5^\circ$ , and added noise with variance 2), and achieves a sub-pixel matching accuracy without increasing the matching time. Compared with the matching methods of fixed rectangular template [20] and the original SIFT-based matching method [21], evaluation results show that the multi-resolution technique shortens the computation time, the method of multi-scale circular template fusion improves the matching performance, and the improved SIFT enables a sub-pixel matching accuracy.

**Table 1.** The statistical results.

Index		Result		Distortion		Zoom out	Zoom in	Rotation	Combined
						(20%)	(20%)	( $7^\circ$ )	distortion
SAR scene matching method	Coarse matching	Matching probability		0.98	0.99	1	0.98		
		Matching error (pixel)		1.9265	1.5335	1.825	2.125		
		Time (s)		0.754	0.821	0.524	0.339		
	Precise matching	Matching error	Position (pixel)	0.3785	0.1825	0.1985	0.4859		
			Angle ( $^\circ$ )	0.101	0.127	0.021	0.152		
		Time (s)		0.279	0.288	0.271	0.261		
Matching by rectangular template		Matching probability		0.90	0.92	0.48	0.25		
		Matching error (pixel)		2.85	2.56	2.39	2.984		
		Time (s)		1.483	1.334	1.728	1.438		
SIFT-based Matching method		Matching probability		0.56	0.59	0.56	0.52		
		Matching error	Position (pixel)	2.92	2.85	2.64	2.862		
			Angle ( $^\circ$ )	0.15	0.105	0.12	0.08		
		Time (s)		5.45	5.52	5.24	5.86		

## 6. CONCLUSION

A fast and robust method for SAR scene matching is presented in this article. Critical techniques include the pyramid multi-resolution decomposition, the multi-scale circular template fusion, the two-stage implementation of image matching, and the improved SIFT-based matching. The pyramid multi-resolution technique and two-stage image matching strategy significantly improves the computational efficiency. The multi-scale circular template fusion and the improved SIFT-based matching prevent geometric distortions from diminishing the accuracy of image matching. The overall position accuracy is achieved at about 0.5 pixels and angle accuracy is  $0.1^\circ$ . This is appropriate for SAR scene matching navigation system.

## REFERENCES

1. Ren, S., W. Chang, T. Jin, and Z. Wang, "Automated SAR reference image preparation for navigation," *Progress In Electromagnetics Research*, Vol. 121, 535–555, 2011.
2. Inglada, J. and F. Adragna, "Automatic multi-sensor image registration by edge matching using genetic algorithms," *IEEE Int. Geosci. Remote Sens. Symposium*, Vol. 5, 2313–2315, Sydney, NSW, Australia, 2001.
3. Liu, H. Z., B. L. Guo, and Z. Z. Feng, "Pseudo-log-polar Fourier transform for image registration," *IEEE Signal Processing Letters*, Vol. 13, No. 1, 17–20, 2006.
4. You, J. and P. A. Bhattacharya, "Wavelet-based coarse-to-fine image matching scheme in a parallel virtual machine environment," *IEEE Trans. Image Processing*, Vol. 9, No. 9, 1547–1559, 2000.
5. Kusiek, A. and J. Mazur, "Analysis of scattering from arbitrary configuration of cylindrical objects using hybrid finite-difference mode-matching method," *Progress In Electromagnetics Research*, Vol. 97, 105–127, 2009.
6. Riabi, M. L., R. Thabet, and M. Belmeguenai, "Rigorous design and efficient optimization of quarter-wave transformers in metallic circular waveguides using the mode-matching method and the genetic algorithm," *Progress In Electromagnetics Research*, Vol. 68, 15–33, 2007.
7. Ren, S., W. Chang, and X. Liu, "SAR image matching method based on improved SIFT for navigation system," *Progress In Electromagnetics Research M*, Vol. 18, 259–269, 2011.
8. Lowe, D. G., "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, 91–110, 2004.
9. Yang, Y., Y. X. Song, F. W. Zhai, et al., "A high-precision localization algorithm by improved SIFT key-points," *2nd Int. Congress on Image and Signal Processing*, 1–6, Tianjin, 2009.
10. Mikolajczyk, K. and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 27, No. 10, 1615–1630, 2005.
11. Ehlers, M., "Multi-sensor image fusion techniques in remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 46, No. 1, 19–30, 1991.
12. Dai, X. L. and S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Trans. Geosci. Remote Sens.*, Vol. 37, No. 5, 2351–2362, 2002.
13. Zhang, Y. C., P. Wang, and G. Y. Zhang, "An information fusion approach and its application based on D-S evidence theory and neural network," *27th Chinese Control Conference Proceedings*, 623–626, Kunming, 2008.
14. Grabner, M., H. Grabner, and H. Bischof, "Fast approximated SIFT," *Asian Conference on Comput. Vis.*, 918–927, Hyderabad, India, 2006.
15. Ke, Y. and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," *Proceedings of IEEE Comput. Soci. Conference on Comput. Vis. Pattern. Recog.*, 506–513, Washington DC, USA, 2004.
16. Yi, Z., Z. G. Cao, and X. Yang, "Multi-spectral remote image registration based on SIFT," *Electronics Letters*, Vol. 44, No. 2, 107–108, 2008.
17. Bastanlar, Y., A. Temizel, and Y. Yardimci, "Improved SIFT matching for image pairs with scale difference," *Electronics Letters*, Vol. 46, No. 5, 346–348, 2010.
18. Xie, H., L. Pierce, and F. Ulaby, "Statistical properties of logarithmically transformed speckle," *IEEE Trans. Geosci. Remote Sens.*, Vol. 40, No. 3, 721–727, 2002.
19. He, F. F., J. Y. Sun, and W. P. Guo, "A practical method for evaluating capability of scene matching algorithms," *7th Int. Conf. Computer-Aided Industrial Design and Conceptual Design*, 1–6, Hangzhou, 2006.
20. Shi, H. L. and B. Hu, "Image registration using a new scheme of wavelet decomposition," *IEEE Instrumentation and Measurement Technology Conference*, 235–239, Victoria, 2008.
21. Li, D. and Y. H. Zhang, "Geometric feature-based image co-registration approach for InSAR," *2nd Asian-Pacific Conf. Synthetic Aperture Radar Proceedings*, 1026–1030, Xi'an, 2009.