Speckle-reducing scale-invariant feature transform match for synthetic aperture radar image registration

Xianmin Wang,[a,b] Bo Li,[a,b] and Qizhi Xu[a,b,*]

[a]Beihang University, School of Computer Science and Engineering, Beijing Key Laboratory of Digital Media, Xueyuan Road No. 37, Haidian District, Beijing 100191, China

[b]Beihang University, State Key Laboratory of Virtual Reality Technology and Systems, Xueyuan Road No. 37, Haidian District, Beijing 100191, China

Abstract. The anisotropic scale space (ASS) is often used to enhance the performance of a scale-invariant feature transform (SIFT) algorithm in the registration of synthetic aperture radar (SAR) images. The existing ASS-based methods usually suffer from unstable keypoints and false matches, since the anisotropic diffusion filtering has limitations in reducing the speckle noise from SAR images while building the ASS image representation. We proposed a speckle reducing SIFT match method to obtain stable keypoints and acquire precise matches for the SAR image registration. First, the keypoints are detected in a speckle reducing anisotropic scale space constructed by the speckle reducing anisotropic diffusion, so that speckle noise is greatly reduced and prominent structures of the images are preserved, consequently the stable keypoints can be derived. Next, the probabilistic relaxation labeling approach is employed to establish the matches of the keypoints then the correct match rate of the keypoints is significantly increased. Experiments conducted on simulated speckled images and real SAR images demonstrate the effectiveness of the proposed method. © *2016 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.10.036030]

Keywords: synthetic aperture radar; image registration; speckle reducing anisotropic diffusion; speckle-reducing anisotropic scale space; probabilistic relaxation labeling.

Paper 16400 received May 26, 2016; accepted for publication Sep. 2, 2016; published online Sep. 26, 2016.

1 Introduction

Synthetic aperture radar (SAR) image registration is one of many key procedures in applications such as change detection, image fusion, and three-dimensional reconstruction.[1] The current registration methods can be generally classified into two categories: area-based methods and feature-based methods. Due to complex geometric deformations and low texture in SAR images, the area-based methods are time consuming and often produce local extrema when estimating the correspondence among the registered images. Compared with the area-based methods, the feature-based methods are computationally efficient and recommended for the SAR image registration, since many distinctive features can usually be obtained from SAR images.[2–4]

The feature-based methods consist of three steps: feature detection, feature description, and feature matching.[5] In the feature detection step, the multiscale space of an image is required to be constructed via filtering the original image with an appropriate function over increasing scale or time. For building the multiscale representation, Gaussian kernel is the simplest option. For instance, the conventional scale-invariant feature transform (SIFT)[6] algorithm utilizes Gaussian filter to blur the image and establish a Gaussian scale space, then the features are detected at different scale levels and resolutions. The conventional SIFT algorithm has been successfully employed to the registration of optical remote sensing images. However, it fails to provide favorable results while dealing with the SAR images. The reason is the fact that Gaussian blurring is one instance of isotropic diffusion filtering which is sensitive to speckle noise and does not

*Address all correspondence to: Qizhi Xu, E-mail: qizhi@buaa.edu.cn

1931-3195/2016/$25.00 © 2016 SPIE

respect the natural boundaries of the object.[7] As a consequence, many unstable keypoints are brought from the Gaussian scale space of SIFT and then the matching performance is degraded. Here, the unstable keypoints are the control points of an image that cannot be repeated in another image being registered.[8]

To enhance the performance of multiscale feature detection of SIFT, it is appropriate to make the blurring adaptive to the image texture so that the noise is smoothed and the image edge remains unaffected. Hence, anisotropic diffusion filters, such as bilateral filter (BF), adapted anisotropic Gaussian (AAG) function, and Perona–Malik equation, are utilized to construct an anisotropic scale space (ASS) for the SAR image registration. The improvements of SIFT algorithm for SAR images have been presented in some literature. Wang et al.[9] proposed a BFSIFT algorithm to obtain accurately located keypoints by generating the ASS with BF. Later, an AAG-SIFT method was introduced for reducing the influence of speckle noises via building the ASS with an AAG filter.[10] Recently, Fan et al.[7] utilized Perona–Malik equation to construct the ASS, where the gradient magnitude of ASS is computed by the ratio of exponential weighted average operator.

These ASS-based methods are able to preserve the image edges and improve the feature location accuracy. However, they cannot effectively reduce the unstable keypoints caused by the speckle noise from SAR images. The reason for this is that in the existing ASS-based approaches, the anisotropic diffusion filters adaptively smooth the noises and preserve the edges due to their different image gradient magnitudes. If the images contain strong multiplicative noises such as speckles, then the image edges are difficult to distinguish from the speckled homogeneous region, since both the image boundaries and the multiplicative noises lead to high image gradient magnitudes.[11] As a result, the speckle noises from the SAR images will be preserved instead of being smoothed by the anisotropic diffusion filters and then identified as unstable keypoints in the ASS.

In the situation of feature matching, the nearest-neighbor distance ratio (NNDR)[6] is the most common method, where the candidate matches are identified based on the nearest neighbor of the keypoints and the best matches are obtained under a distance ratio. For further improving the effectiveness of NNDR in the SAR image registration, Wang et al.[9] proposed a dual matching (DM) method in which the NNDR strategy is adopted in the reference and sensed image separately. By analyzing the dominant orientation consistency (DOC) properties of the keypoints, a DOC method is introduced to enhance the keypoint correct match rate for SAR images. These methods establish the keypoint matches based on the maxima similarity of the keypoint descriptor. Due to similar patterns in SAR images, a pair of keypoints associated with the maximum similarity does not always correspond to the best match pair. Moreover, we observe that the correlation of the keypoints is a multipeak function, then the similarity of the correct matching may not be at the maximum peak. Therefore, it is reasonable to identify the matches based on the keypoint global consistency.

In this paper, we proposed a speckle reducing SIFT (SR-SIFT) match method to obtain stable keypoints and precise matches for the SAR image registration. The contributions of this paper are as follows. First, a speckle reducing anisotropic scale space (SRASS) is constructed based on the speckle reducing anisotropic diffusion (SRAD). Due to the gradient magnitude operator and the Laplacian operator of SRAD, speckle noises are greatly reduced and the edges of the images are preserved in the SRASS, then the stable keypoints can be obtained. Second, we utilize the probabilistic relaxation labeling (PRL) approach to establish the matches of the keypoints via a global optimization process, where the compatibility coefficient of PRL is evaluated depending upon the adaptive support windows. With the PRL matching, the keypoint correct match rate is significantly enhanced. We validate our method on simulated images and real SAR images and the experimental results demonstrate the effectiveness of our method.

2 Proposed Method

The proposed SR-SIFT match algorithm focuses on two improvements: (1) utilizing the SRAD to construct an SRASS for obtaining stable keypoints and (2) employing PRL matching approach for increasing the correct match rate of keypoints. The main steps of the proposed method are shown in Fig. 1.

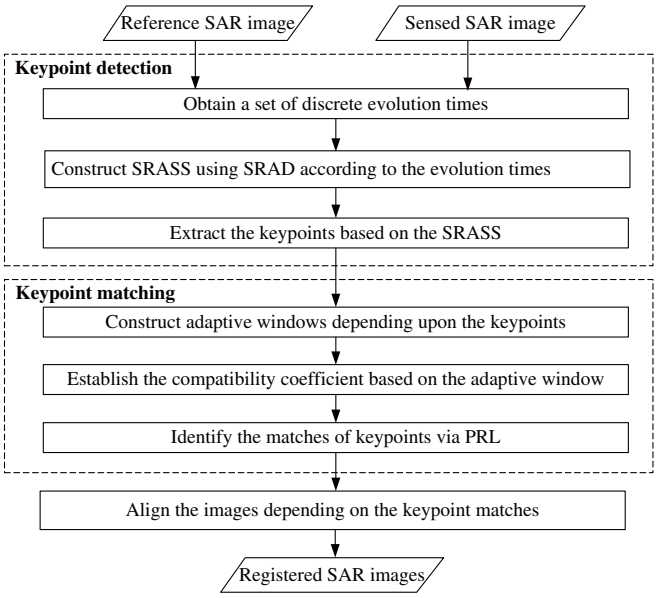

Fig. 1 The main steps of the proposed method.

2.1 *Speckle Reducing Anisotropic Scale Space*

The existing ASS-based methods overcome the shortcomings of the conventional SIFT algorithm based on anisotropic diffusion filtering. However, they suffer from unstable keypoints caused by speckle noises in SAR images, since the anisotropic diffusion filters detect edges depending upon image gradient magnitude and would not smooth the speckled homogeneous regions. To enhance the stability of the keypoint detection, we construct an SRASS using the SRAD. Then the keypoints are detected in the SRASS.

The SRAD[12] is an edge-sensitive partial differential equation version of the conventional speckle reducing filters, such as Lee and Frost. Let $I_0(x,y)$ be the origin intensity image and $I(x,y;t)$ be the filtered image, then SRAD can be written as

$$\begin{cases} \partial I(x,y;t)/\partial t = \mathrm{div}[c(q)\cdot\nabla I(x,y;t)] \\ I(x,y;0) = I_0(x,y) \end{cases}, \tag{1}$$

where div and ∇ are divergence and gradient operators, and the time t is the scale parameter. In Eq. (1), $c(q)$ refers to the conductivity coefficient defined as

$$c(q) = \frac{1}{1+[q^2(x,y;t)-q_0^2(t)]/\{q_0^2(t)[1+q_0^2(t)]\}}. \tag{2}$$

The conductivity coefficient $c(q)$ is a monotone increasing function in scope (0, 1), which controls the diffusion process. Small $c(q)$ leads to slow diffusion speed thus the image texture is preserved. While large $c(q)$ results in fast diffusion speed, thus the image texture is smoothed. The argument $q(x,y;t)$ in the conductivity coefficient serves as an edge detector for SRAD determined by

$$q(x,y;t) = \sqrt{\frac{(1/2)(|\nabla|/I)^2-(1/16)(\nabla^2 I/I)^2}{[1+(1/4)(\nabla^2 I/I)]^2}}. \tag{3}$$

The edge detector $q(x,y;t)$ exhibits different values for the image edges and the speckled regions on the basis of the normalized gradient magnitude operator $|\nabla|/I$ and the normalized Laplacian operator $\nabla^2 I/I$. At the center of an edge, the Laplacian term undergoes zero crossing and the gradient term dominates, leading to a relatively large $q(x,y;t)$. Then the conductivity

coefficient approaches 0 and the edge is preserved. While in the speckled homogeneous regions, the normalized image divergence is approximately equal to the normalized gradient magnitude, resulting in a relatively small $q(x, y; t)$. Thus the conductivity coefficient closes to 1 and the speckle noise is smoothed. The $q_0(t)$ is the diffusion threshold estimated by $q_0(t) = \sqrt{\text{var}[z(t)]}/\overline{z(t)}$, where $\text{var}[z(t)]$ and $\overline{z(t)}$ are the variance and mean of the image intensity over a homogeneous area at the time t. Additionally, $q_0(t)$ controls the amount of smoothing applied to the image by SRAD.

In applications, since there are no analytical solutions for Eq. (1), a numerical method is used to make an approximation of SRAD. In Ref. 12, the SRAD is implemented based on the simplest finite difference discretization by means of a so called explicit or Euler-forward scheme. This scheme requires very small time steps in order to be stable, which leads to poor efficiency and limits its applications. Fortunately, Weickert et al.[13] presented a semi-implicit schema to enhance the effectiveness and efficiency of solving the anisotropic diffusion filtering equation in Ref. 13. The backbone of this schema is the use of the additive operator splitting (AOS) technique, which is rotationally invariant and computationally efficient. Moreover, the semi-implicit schema is stable for all time steps. By means of the semi-implicit schema, Eq. (1) can be discretized and reconstructed as an iterative form

$$I^{i+1} = \left[E - (t_{i+1} - t_i)\sum_{l=1}^{m} A_l(I^i)\right]^{-1} I^i, \tag{4}$$

where I^i and I^{i+1} are the image vector representations at time t_i and t_{i+1}, E is the identity matrix, l denotes the discretizing direction, and $A_l(I^i)$ is a matrix which encodes the conductivity coefficient $c(q)$ along the l'th dimension. Suppose the image has M rows and N columns, the process of image smoothing can be separated into two one-dimensional cases based on the AOS strategy. In the case of filtering the image along the horizontal dimension (i.e., $l = x$, along the x coordinate axis), let I_x^{i+1} be the image at the time t_{i+1} and $I_{x,r}^{i+1}$ be the r'th row of I_x^{i+1}. Then $I_{x,r}^{i+1}$ corresponds to the solution of equation $[E - 2(t_{i+1} - t_i)A_x(I^i)]I_{x,r}^{i+1} = I_{x,r}^i$, in which $I_{x,r}^{i+1}$ is an $N \times 1$ vector and $A_x(I^i)$ is an $N \times N$ matrix whose element a_{uv} is determined by

$$a_{uv} = \begin{cases} (c_u^i + c_v^i)/2 & v \in N(u) \\ -\sum\limits_{k \in N(u)} (c_u^i + c_k^i)/2 & v = u \\ 0 & \text{else} \end{cases}, \tag{5}$$

where u is the index of the pixel in vector $I_{x,r}^{i+1}$, $N(u)$ represents the neighborhood of u, and c_u^i is the conductivity coefficient of the pixel associated with u at the time t_i depending on Eq. (2). Hence, the image I_x^{i+1} along the horizontal dimension can be derived row by row according to $I_{x,r}^{i+1}$ for $r = 1, \ldots, M$. Meanwhile, the image I_y^{i+1} at the time t_{i+1} along the vertical dimension (i.e., $l = y$, along the y coordinate axis) can be obtained in the same manner. As a consequence, the final blurred image I^{i+1} at the time t_{i+1} can be determined by $I^{i+1} = (I_x^{i+1} + I_y^{i+1})/2$. More details about how to construct $A_l(I^i)$ and compute I^{i+1} can be found in Ref. 13. Note that since the system matrix $A_l(I^i)$ is tridiagonal and diagonal dominant, Eq. (4) can be efficiently solved by the Thomas algorithm.

For constructing the multiscale image representation, we divide the scale space of the image into O octaves and S sublevels logically as done in SIFT. Note that we always perform the blurring at the resolution of the origin image, instead of downsampling the image at each new octave. Suppose that the set of octaves and sublevels are identified by index o and index s, then o and s can be mapped to their corresponding scale by

$$\sigma_i(o, s) = \sigma_0 2^{o+s/S}, \quad o \in [0, O-1], \quad s \in [0, S-1], \quad i \in [0, N], \tag{6}$$

where σ_0 is the base scale and $N = O \times S$ is the total number of blurred images in the multiscale space. Since the anisotropic diffusion filtering is defined in time terms, the discrete scale σ_i is required to be converted into time units using the equation[14]

$$t_i = \sigma_i^2/2. \tag{7}$$

From Eq. (7), we obtain a set of discrete evolution times $\{t_0, t_1, \ldots t_N\}$ for $i = 0, \ldots, N$. Then, according to the evolution times $\{t_0, t_1, \ldots t_N\}$ and the matrix $A_l(I^i)$, it is straightforward to derive a set of blurred images $\{I^1, I^2, \ldots, I^N\}$ in an iterative way using Eq. (4). Thus, the SRASS image representation L_{SRASS} can be expressed as

$$L_{\mathrm{SRASS}} = \{I^1(t_1), I^2(t_2), \ldots, I^N(t_N)\}, \tag{8}$$

where t_i is the scale level in the time term corresponding to the blurred image I^i in the SRASS ($i = 1, \ldots, N$). In our experiments, the base scale σ_0 and the total number of octaves O in Eq. (6) are set at 1.6 and 3 as suggested in Ref. 6. Moreover, our experimental results indicate that the performance of the proposed method is robust when the total number of the blurred images $N \in [8,10]$. If $N > 10$, it is computationally expensive for constructing the SRASS representation while $N < 8$, and we cannot obtain a reasonable quantity of stable keypoints for accurately evaluating the transform parameters. Thus, an appropriate total number of the blurred images N is 9 and the quantity of sublevels S in each octave is 3. Figure 2 shows a comparison between the SRASS and the anisotropic Gaussian space (AGSS) for several evolution times given the same reference SAR image. As can be observed, strong speckle noises are smoothed and prominent structures are preserved in the SRASS, whereas in the AGSS the speckle noises are retained or even enhanced.

After the SRASS have been constructed, the difference between adjacent smoothed images in the SRASS is performed. Then the keypoints are detected as done in the SIFT algorithm. In the SRASS, the speckle noises are significantly reduced and the prominent structure of the images is preserved, thus stable and accurately located keypoints can be obtained.

2.2 *Probabilistic Relaxation Labeling Matching Approach with Adaptive Support Window*

The PRL is an iterative process that propagates label probabilities globally via local interactions.[15] Suppose that the keypoint sets of the reference and sensed image are $P = \{p_i\}_1^m$

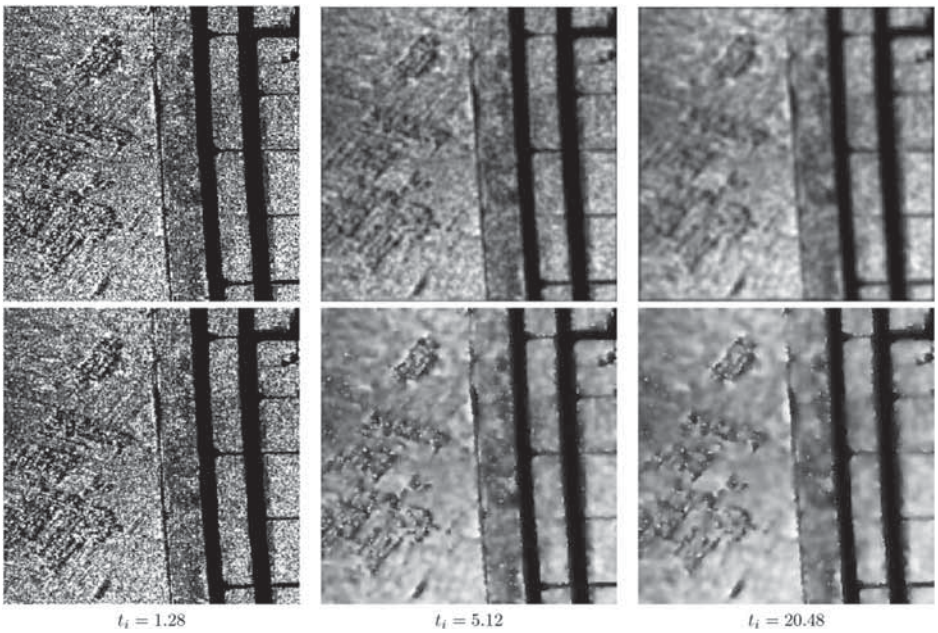

Fig. 2 Comparison between the SRASS and anisotropic Gaussian scale space (AGSS) for several evolution times t_i. The first row is the AGSS which is built by anisotropic Gaussian filtering, where the image ENL values at the time $t_i = 1.28$, 5.12, and 20.48 are 1.12, 2.42, and 3.53, respectively. The second row is the SRASS constructed based on SRAD, where the image ENL values at the time $t_i = 1.28$, 5.12, and 20.48 are 1.24, 3.73, and 6.56, respectively.

and $Q = \{q_j\}_1^n$, the keypoint matching can be regarded as a task of assigning a consistent unambiguous label $q_j \in Q$ to each object $p_i \in P$ on the basis of contextual information. In the iterative process, the compatibility coefficient plays an important role in updating the label probabilities. Since the keypoints produce similar descriptors in SAR images, the adaptive support window relying upon the keypoints is employed for establishing the keypoint local interactions. Then the normalized cross-correlation (NCC)[16] measure of the corresponding windows in the reference image and the sensed image is used as the compatibility coefficient of PRL. Unlike the conventional fixed size rectangle windows, the support window in our method is referred to as the image line segment whose endpoints are the keypoints of the image. Thus, the proposed window can be adaptively constructed depending on the distribution of the keypoints, where the size of the support window is determined by the distance of the chosen keypoints. Moreover, if the corresponding support windows in the reference image and sensed image are in different sizes, a resampling should be performed to ensure they have the same size before comparing their similarities. By this method, the impact of image geometric distortion is significantly reduced and the matching performance can be improved. The PRL matching method is carried out as below.

As shown in Fig. 3, let p_h and q_k be any keypoints other than p_i and q_j in P and Q, and the $\delta_{ij}(h,k)$ be the compatibility coefficient between p_h and q_k when p_i pairs with q_j. Then the image line segment from p_i to p_h is defined as the target support window L_P, and the image line segment from q_j to q_k is defined as the search support window L_Q. Since the length of L_P is different from the length of L_Q, we resample L_P and L_Q to the same size. Note that the compatibility coefficient $\delta_{ij}(h,k)$ based on the NCC ranges from -1 to 1. If $\delta_{ij}(h,k)$ is equal to 1 (or -1), the pairs (p_h, q_k) and (p_i, q_j) have the maximal (or minimal) compatibility, so that (p_h, q_k) should provide (p_i, q_j) with the maximal (or minimal) support. Meanwhile, the support of (p_i, q_j) declines with the decrease of $\delta_{ij}(h,k)$. Because the range of the support function is in the scope [0, 1], the support of (p_i, q_j) provided by (p_h, q_k) can be represented as

$$\phi_{ij}(h,k) = 1/\left\{1 + e^{-\tan[\frac{\pi}{2}\delta_{ij}(h,k)]}\right\}. \tag{9}$$

For each p_h, it is required that only one q_k corresponds to p_h when p_i is paired with q_j. Hence, the support of (p_i, q_j) associated with p_h is defined as $\max_{k\neq j}[\phi_{ij}(h,k)]$. To obtain the total support $s_{ij}^{(0)}$ for (p_i, q_j), we average the contributions of all p_h

$$s_{ij}^{(0)} = \frac{1}{m-1}\sum_{h\neq i}\max_{k\neq j}[\phi_{ij}(h,k)]. \tag{10}$$

In computing $s_{ij}^{(0)}$, any two keypoints can be paired at the initial iteration. While at the r'th iteration $(r > 0)$, the support of (p_i, q_j) given by (p_h, q_k) should depend on $\phi_{ij}(h,k)$ as well as its previous support $s_{hk}^{(r-1)}$, which allows local support to propagate. Therefore, we take the average of theses and the support function $s_{ij}^{(r)}$ at the r'th iteration can be defined as

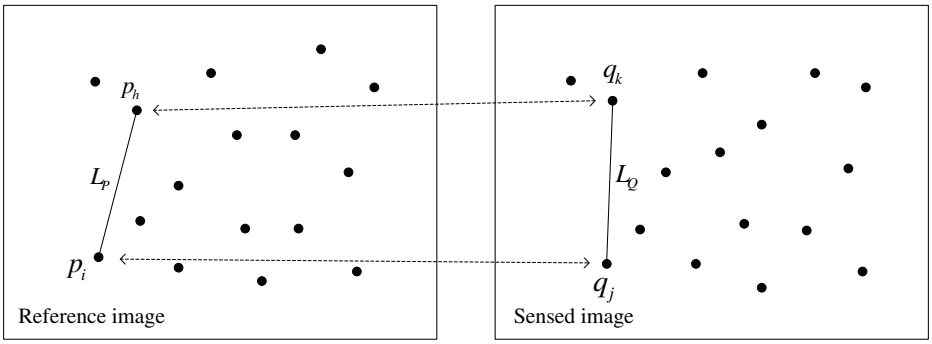

Fig. 3 The construction of adaptive support windows for PRL, where p_i is brought into pairing with q_j, and p_h is brought into pairing with q_k. L_P is the target window in the reference image and L_Q is the search window in the sensed image.

$$s_{ij}^{(r)} = \frac{1}{2(m-1)} \sum_{h \neq i} \max_{k \neq j} [s_{hk}^{(r-1)} + \phi_{ij}(h,k)]. \tag{11}$$

Let $S^{(0)} = [s_{ij}^{(0)}]_{m \times n}$ and $S^{(r)} = [s_{ij}^{(r)}]_{m \times n}$ be the matching probability matrices at the initial and r'th iteration. In the iteration process, the matching probabilities of correct matches are increasing, while the matching probabilities of false matches are declining. Moreover, the iteration process stops if either the maximum number of the iteration is achieved or the absolute error of matching probabilities at the r'th and $(r-1)$'th iteration is smaller than a given threshold ε, i.e.,

$$\delta(r) = \sum_{i=1}^{m} \sum_{j=1}^{n} [s_{ij}^{(r)} - s_{ij}^{(r-1)}] < \varepsilon. \tag{12}$$

Let $S^{(l)}$ be the matching probability matrix at the last iteration. (p_i, q_j) is identified as a pair of matched keypoints if $s_{ij}^{(l)}$ is the maximum element of $S^{(l)}$ in that row and column.

3 Experimental Results

To validate the effectiveness of the proposed method, we conduct experiments on a sequence of simulated speckled images and pairs of real SAR images taken from different times, polarizations, and incidence angles. We first investigate the stability and robustness of keypoint detection on the sequence of simulated images. Then we evaluate the keypoint matching performance using the datasets of real SAR images. All the experiments are carried out on a workstation with Intel Xeon CPU and 16 GB memory.

3.1 *Stability and Robustness of Keypoint Detection*

In this section, the stability and robustness of the keypoint detectors are tested with the help of the repeatability rate. The repeatability rate evaluates the ratio of the number of keypoints repeated between two images to the total number of detected keypoints under a given location error ϵ.[8] The repeatability criterion provides a measure of the stability of keypoint detection, where a high repeatability rate presents strong stability of the keypoint descriptor. For creating the sequence of simulated speckled images, we add multiplicative noise to an original gray image based on the multiplicative model. Since the observed intensity at each resolution cell of SAR images can be formulated by a multiplicative noise of unit mean,[17] we construct the simulated image sequence using the equation $L_i = H_i \times L_0$ $(i = 1, \ldots, 10)$, where L_0 is the origin image, L_i is the i'th speckled image in the sequence, H_i is the Gaussian noise of mean 1, and the standard deviation is $0.1 \times i$. The equivalent number of look (ENL) values of the images in the sequence can be evaluated by $(0.1 \times i)^{-1}$,[18] for $i = 1, \ldots, 10$. Figure 4 shows a part of the simulated sequence.

To test the keypoint detection performance against speckle noise, the repeatability rates between L_0 and L_i $(i = 1, \ldots, 10)$ are evaluated. We compare our method with several ASS-based algorithms designed for SAR image registration, including BFSIFT,[9] AAG-SIFT,[10] and NDSS-SIFT (NDSS denotes nonlinear diffusion scale space).[7] We also compare the proposed method with SAR-SIFT,[19] whose main modification is based on the SAR-Harris scale space constructed by a new multiscale SAR-Harris function. The repeatability rates of the five methods are shown in Fig. 5(a). It can be observed that although the repeatability rates of the five methods decline with the variance of speckle noises, SR-SIFT always exhibits a better stability in terms of keypoint detection. Note that the keypoint detection performance of SAR-SIFT is unstable. It obtains comparable repeatability rates with SR-SIFT in some cases, while getting low repeatability rates in other cases. It is also noticeable that the repeatability rate obtained by SR-SIFT falls more slowly than the repeatability rate acquired using the other four methods. These facts imply that SR-SIFT produces a favorable result for reducing the impact of speckle noises. As a consequence, unstable keypoints caused by the noises can be significantly suppressed.

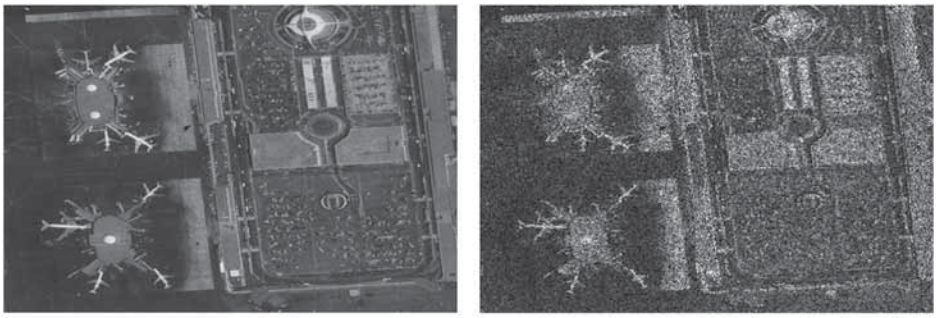

Fig. 4 A part of the simulated image sequence. (a) The origin gray image L_0. (b) The speckled image L_6 with noise variance 0.6, where the ENL value is 1.67.

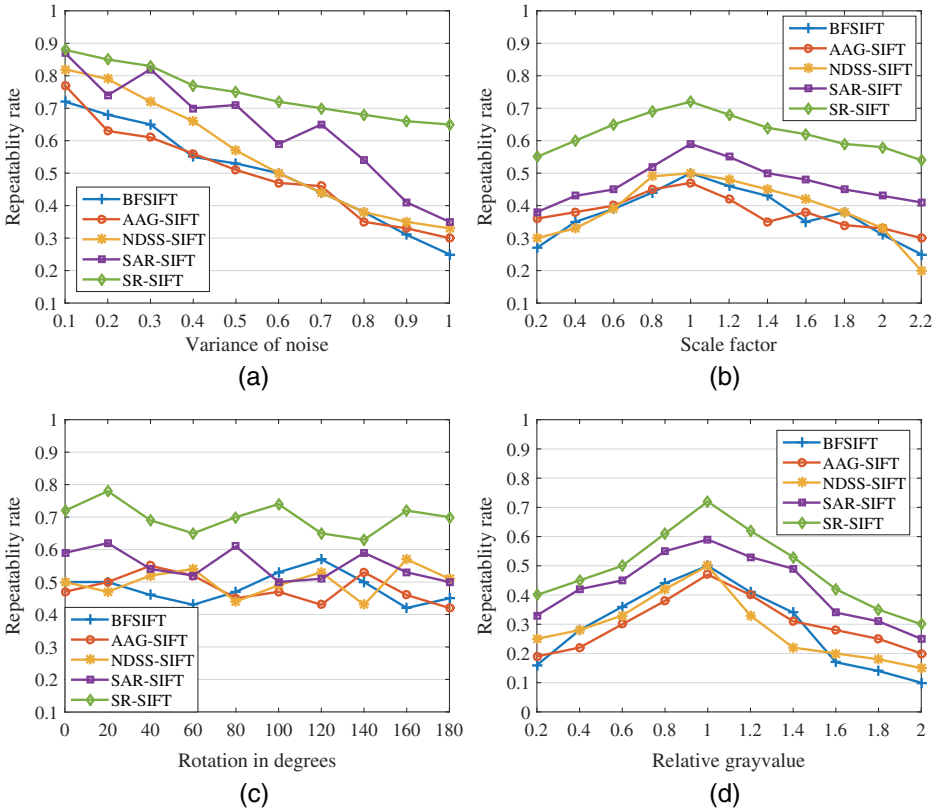

Fig. 5 Repeatability rates of the five testing methods, where the location error $\epsilon = 1.2$ pixels. (a) Repeatability rate for speckle noise variance. (b) Repeatability rate for scale change. (c) Repeatability rate for image rotation. (d) Repeatability rate for uniform illumination variation.

Furthermore, the repeatability rates are evaluated between L_0 and L_6 under various transformations including scale change, image rotation, and illumination variation. The experimental results are shown in Figs. 5(b)–5(d). It can be observed that the repeatability rates of SR-SIFT outperform its competitors for all the analyzed sequences. This fact demonstrates that SR-SIFT is a more robust local invariant feature detector for the registration of speckled images.

3.2 *Performance of Keypoint Matching*

Six datasets consisting of various SAR images are used in our second experiment.

1. Dataset 1: two SAR images taken over an agricultural scene from ALOS-PALSAR system with a size of 1146×1250 pixels are included in dataset 1. The reference image [left image in Fig. 6(a)] is obtained for the VV polarization in May, 2007, and the sensed image [right image in Fig. 6(a)] is obtained for the HV polarization in March, 2009. The ENL values of the reference image and sensed image are 8.03 and 8.12.
2. Dataset 2: dataset 2 contains two SAR images collected by TerraSAR-X satellite with a size of 1035×1024 pixels in the area of Toulouse, France, at different incidence angles. The reference image [left image in Fig. 6(c)] is taken at an incidence angle of 34 deg and the sensed one [right image in Fig. 6(c)] is taken at an incidence angle of 29 deg. The ENL values of the reference image and sensed image are 3.01 and 2.97.
3. Dataset 3: this dataset includes two SAR images obtained from NASA/JPL AIRSAR, an airborne SAR sensor used by the National Aeronautics and Space Administration/Jet Propulsion Laboratory. The reference image [left image in Fig. 6(e)] has a size of 1102×1136 pixels and the sensed image [right image in Fig. 6(e)] has a size of 1346×1398 pixels. The ENL values of the reference image and sensed image are 1.15 and 1.08. Note that there exist substantial relative rotations between the two images in dataset 3.
4. Dataset 4: this dataset consists of two 1324×1320 pixels SAR images taken by Radarsat-2 covering the area of Yellow River Estuary. The reference image [left image in Fig. 6(g)] is single-look SAR image acquired in 2009, and the sensed image [right image in Fig. 6(g)] is fourth-look SAR image obtained in 2008. The ENL values of the reference image and sensed image are 1.12 and 5.21. Note that the two images in dataset 4 have very different noise levels.
5. Dataset 5: dataset 5 contains two SAR images sized of 1143×1121 pixels with resolution of 3 m. The reference image [left image in Fig. 6(i)] is taken by Radarsat-2 and the sensed one [right image in Fig. 6(i)] is taken by TerraSAR. The ENL values of the reference image and sensed image are 6.16 and 7.43. Note that there has been a great difference between the reference and sensed images in terms of texture, since the two images are coming from different sensors.
6. Dataset 6: two airborne X- and P-band SAR images taken from different look angles are included in dataset 6. These images are captured from a rural area with a size of 1426×1420 pixels. The reference image [left image in Fig. 6(k)] is X-band image obtained on December 7, 2010, and the sensed image [right image in Fig. 6(k)] is P-band image acquired on December 4, 2010. The ENL values of the reference image and sensed image are 7.21 and 7.28.

In the experiments, we first obtain the keypoints by the proposed method (the repeatability rates of the proposed method for dataset 1 to 6 are 0.88, 0.76, 0.67, 0.61, 0.70, and 0.72, respectively), and then identify the matches of keypoints by NNRD, DM, DOC, and PRL matching methods. The experimental results of the four matching methods for the six datasets are reported in Tables 1–6, where CMN is the number of correct matches, TMN is the total number of keypoint matches, CMR is the correct match rate that is determined by the ratio of CMN to TMR, and RMSE is the root-mean-square error evaluated by the method in Ref. 20. From Tables 1–6, we can see that PRL always has the greatest value of CMN and CMR in the four comparing methods for the six datasets. The reason is that in our matching method, the keypoints with similar descriptors can be effectively differentiated via the global optimization process. Meanwhile, the adaptive support windows constructed based on the keypoints also improve the matching stability against image geometric distortions. In addition, our method exhibits a higher registration accuracy according to the RMSE values in Tables 1–6, which indicates that a greater number of correct matches contributes to calculating accurate transformation parameters.

The iteration processes of PRL matching for the testing datasets are shown in Fig. 6. It can be seen that although many false matches appear at the initial iteration for the six datasets [see Figs. 6(a), 6(c), 6(e), 6(g), 6(i), and 6(k)], the false matches are significantly filtered out by PRL matching in the global optimization process at the last iteration [see Figs. 6(b), 6(d), 6(f), 6(h), 6(j), and 6(l)]. This phenomenon can be explained by Fig. 7. As shown in Fig. 7, the CMR of PRL increases with the increase of the iteration time and the iteration process stops when the CMR reaches steady values. In addition, Fig. 7 shows that the convergence speed of PRL matching for dataset 1 is faster than that for the other five datasets. The reason is the fact that the images

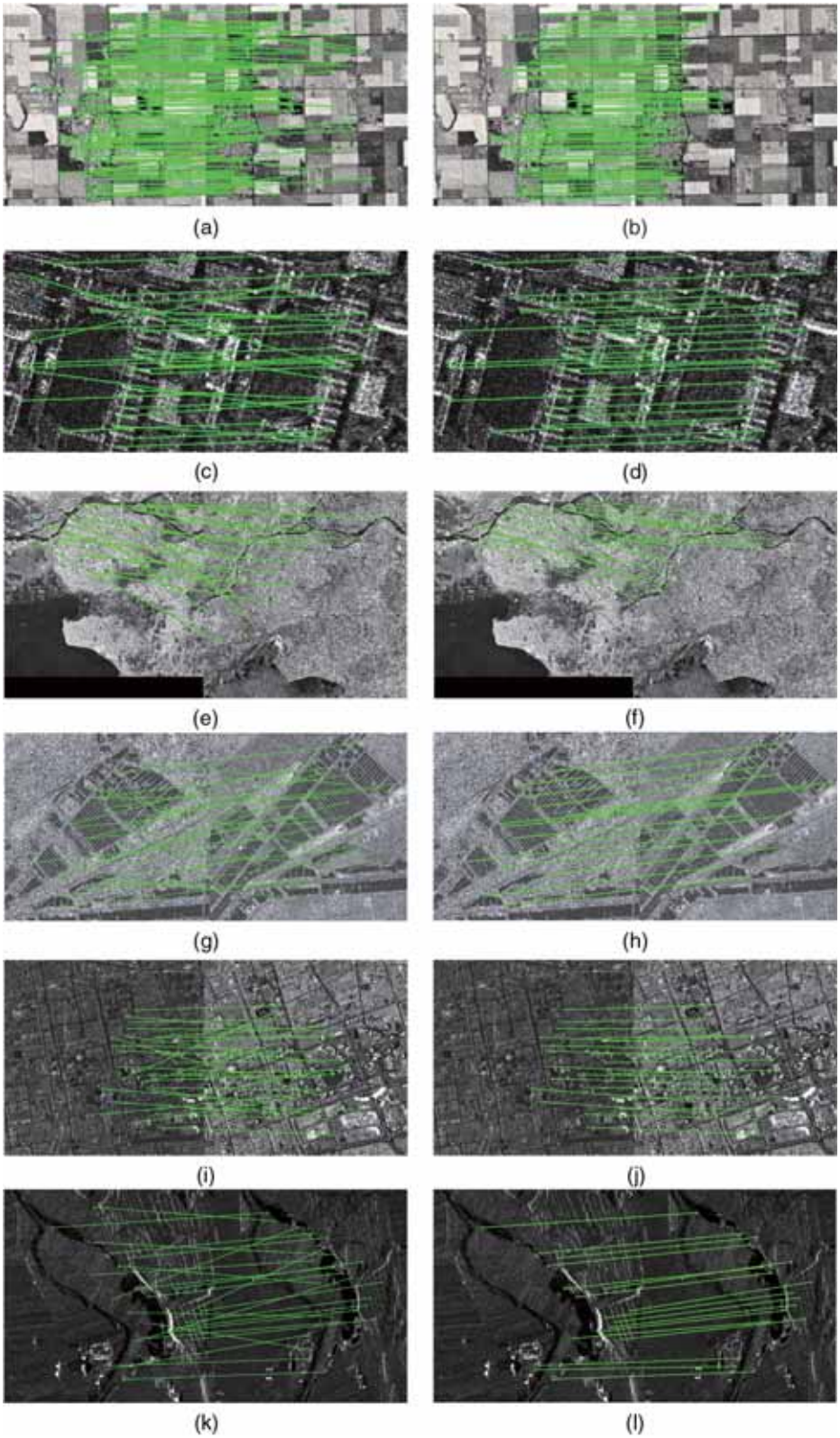

Fig. 6 The iteration processes of PRL matching for datasets 1–6. (a), (c), (e), (g), (i), and (k) The matching results at the initial iteration for datasets 1–6. (b), (d), (f), (h), (j), and (l) The matching results at the last iteration for datasets 1–6.

Table 1 Matching performance for dataset 1.

	NNDR	DM	DOC	PRL
CMN	212	232	254	323
TMN	408	380	339	344
CMR	0.52	0.61	0.75	0.94
RMSE	0.94	0.73	0.56	0.41

Table 2 Matching performance for dataset 2.

	NNDR	DM	DOC	PRL
CMN	128	142	165	198
TMN	267	245	299	225
CMR	0.48	0.58	0.72	0.88
RMSE	1.28	0.96	0.82	0.51

Table 3 Matching performance for dataset 3.

	NNDR	DM	DOC	PRL
CMN	152	176	197	231
TMN	287	275	281	272
CMR	0.53	0.64	0.70	0.85
RMSE	1.12	0.83	0.74	0.55

Table 4 Matching performance for dataset 4.

	NNDR	DM	DOC	PRL
CMN	171	185	212	265
TMN	305	298	290	308
CMR	0.56	0.62	0.73	0.86
RMSE	1.09	0.80	0.72	0.51

Table 5 Matching performance for dataset 5.

	NNDR	DM	DOC	PRL
CMN	127	162	183	201
TMN	254	261	251	245
CMR	0.50	0.62	0.73	0.82
RMSE	1.21	0.83	0.77	0.68

Table 6 Matching performance for dataset 6.

	NNDR	DM	DOC	PRL
CMN	112	148	156	176
TMN	238	239	216	217
CMR	0.47	0.62	0.72	0.81
RMSE	1.56	0.92	0.83	0.75

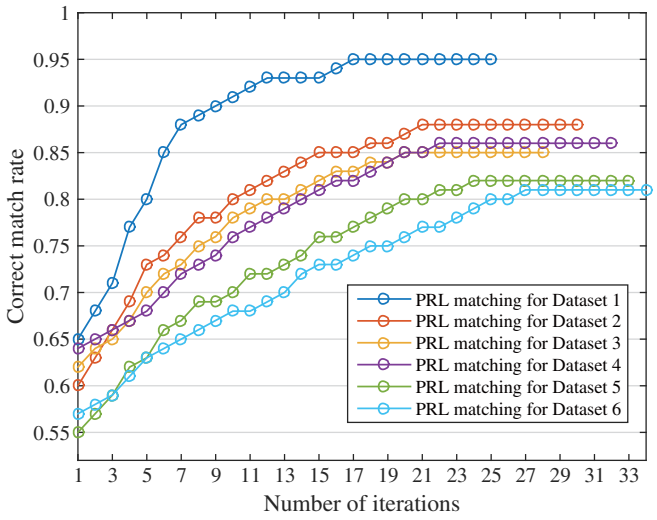

Fig. 7 The plot of CMR varying with iteration times for datasets 1–6 in the PRL matching.

in dataset 1 contain more detailed information of texture, providing stronger support for the correct matches in the global iteration process.

4 Conclusions

This paper presents an SR-SIFT method to obtain stable keypoints and precise matches for the registration of SAR images. To reduce the unstable keypoints caused by the speckle noise, an SRASS based on the SRAD is constructed. Moreover, to enhance the correct match rate, the PRL matching is employed to establish the keypoint matches in an iterative way. Experiments are carried on a sequence of simulated speckled images and pairs of real SAR images. The experimental results demonstrate the effectiveness of the proposed method.

Acknowledgments

This work was partially supported by the National Science Fund for Distinguished Young Scholars (No. 61125206) and the NSFC Key Project (No. 61331017).

References

1. B. Zitova and J. Flusser, "Image registration methods: a survey," *Image Vision Comput.* **21**(11), 977–1000 (2003).
2. W. Ren et al., "Registration of multitemporal low-resolution synthetic aperture radar images based on a new similarity measure," *J. Appl. Remote Sens.* **10**(1), 015001 (2016).
3. L. Yang et al., "Robust image registration using adaptive coherent point drift method," *J. Appl. Remote Sens.* **10**(2), 025014 (2016).

4. J. Jiang, S. Zhang, and S. Cao, "Rotation and scale invariant shape context registration for remote sensing images with background variations," *J. Appl. Remote Sens.* **9**(1), 095092 (2015).
5. Q. Xu, Y. Zhang, and B. Li, "Improved sift match for optical satellite images registration by size classification of blob-like structures," *Remote Sens. Lett.* **5**(5), 451–460 (2014).
6. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision* **60**(2), 91–110 (2004).
7. J. Fan et al., "SAR image registration using phase congruency and nonlinear diffusion-based sift," *IEEE Geosci. Remote Sens. Lett.* **12**(3), 562–566 (2015).
8. C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *Int. J. Comput. Vision* **37**(2), 151–172 (2000).
9. S. Wang, H. You, and K. Fu, "BFSIFT: a novel method to find feature matches for SAR image registration," *IEEE Geosci. Remote Sens. Lett.* **9**(4), 649–653 (2012).
10. F. Wang, H. You, and X. Fu, "Adapted anisotropic Gaussian sift matching strategy for SAR registration," *IEEE Geosci. Remote Sens. Lett.* **12**(1), 160–164 (2015).
11. J. Weickert, "A review of nonlinear diffusion filtering," *Lect. Notes Comput. Sci.* **1252**, 3–28 (1997).
12. Y. Yu and S. T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Trans. Image Process.* **11**(11), 1260–1270 (2002).
13. J. Weickert et al., "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans. Image Process.* **7**(3), 398–410 (1998).
14. P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "Kaze features," in *Computer Vision (ECCV 2012)*, pp. 214–227, Springer (2012).
15. H. Wang and E. R. Hancock, "Probabilistic relaxation labelling using the Fokker–Planck equation," *Pattern Recognit.* **41**(11), 3393–3411 (2008).
16. F. Zhao, Q. Huang, and W. Gao, "Image matching by normalized cross-correlation," in *2006 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2006) Proc.*, Vol. **2**, IEEE (2006).
17. L. M. Kaplan, "Analysis of multiplicative speckle models for template-based SAR ATR," *IEEE Trans. Aerosp. Electron. Syst.* **37**(4), 1424–1432 (2001).
18. J. J. Ranjani and S. Thiruvengadam, "Generalized SAR despeckling based on DTCWT exploiting interscale and intrascale dependences," *IEEE Geosci. Remote Sens. Lett.* **8**(3), 552–556 (2011).
19. F. Dellinger et al., "SAR-sift: a sift-like algorithm for SAR images," *IEEE Trans. Geosci. Remote Sens.* **53**(1), 453–466 (2015).
20. X. Wang et al., "Robust and fast scale-invariance feature transform match of large-size multispectral image based on keypoint classification," *J. Appl. Remote Sens.* **9**(1), 096028 (2015).

Xianmin Wang received his BS degree from Suzhou University, Jiangsu, China, in 2006, and his MS degree in computer science from Jiangxi University of Science and Technology, Jiang Xi, China, in 2013. Currently, he is working toward his PhD in computer science from Beihang University, Beijing, China. His research interests include image registration, information fusion, computer vision, and image understanding.

Bo Li received his BS degree from Chongqing University, Shapingba, China, in 1986, his MS degree from Xian Jiaotong University, Xian, China, in 1989, and his PhD from Beihang University, Beijing, China, in 1993, all in computer science. Currently, he is a professor of computer science and engineering and the director of the Beijing Key Laboratory of Digital Media with Beihang University.

Qizhi Xu received his BS degree in computer science from Jiangxi Normal University, Jiangxi, China, in 2005, and his PhD in computer science from Beihang University, Beijing, China, in 2012. He is a postdoctoral fellow with the Canada Research Chair Laboratory in Advanced Geomatics Image Processing, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, New Brunswick, Canada. His research interests include image registration, information fusion, computer vision, and image understanding.