On the Appropriate Feature for General SAR Image Registration

Dong Li*, Yunhua Zhang
Key Lab of Microwave Remote Sensing, Center for Space Science and Applied Research, Chinese Academy of Sciences, No.1 Nanertiao, Zhongguancun, Haidian District, Beijing 100190, China

ABSTRACT

An investigation to the appropriate feature for SAR image registration is conducted. The commonly-used features such as tie points, Harris corner, the scale invariant feature transform (SIFT), and the speeded up robust feature (SURF) are comprehensively evaluated in terms of several criteria such as the geometrical invariance of feature, the extraction speed, the localization accuracy, the geometrical invariance of descriptor, the matching speed, the robustness to decorrelation, and the flexibility to image speckling. It is shown that SURF outperforms others. It is particularly indicated that SURF has good flexibility to image speckling because the Fast-Hessian detector of SURF has a potential relation with the refined Lee filter. It is recommended to perform SURF on the oversampled image with unaltered sampling step so as to improve subpixel registration accuracy and speckle immunity. Thus SURF is more appropriate and competent for general SAR image registration.

Keywords: Feature detector, feature descriptor, image registration, speeded up robust feature (SURF), subpixel accuracy, synthetic aperture radar (SAR).

1. INTRODUCTION

Benefit from a number of SAR missions, the available SAR imagery of one region increases dramatically, which makes the joint processing of multiple images to accurately sense and understand the scene possible, and this may become a potential direction for SAR image processing. Since SAR images may be acquired from different imaging geometries and/or by different sensors, there is always a geometrical warp between the images and this should be aligned first to enable further application. The task of image registration is to estimate the warp function between images so that the same pixel position in each image is mapped to the same target position in the global coordinate system.

Lots of SAR image registration algorithms have been hitherto proposed, and in this paper we only concentrate on the algorithms which conduct the registration based on image feature. The contour [1], [2], region [3], [4], line [5], [6], and point [7], [8] are the commonly-used features for SAR image registration. The first three features and their combination, i.e. the multi-features [9], [10] are usually used for the registration of images from different modalities, such as SAR and optical camera. For the registration of SAR images with inherent speckle and distortion, point feature is much clearer and easier extraction. The physical point, tie points, corner, and keypoint are the commonly-used features for SAR image registration. The physical point refers to the distinctive existing target in the real world, such as the road junction and crossroad [11], [12], building corner [13], ground control point [14]-[16], isolated point scatterer [17], and even the temporarily as well as partially coherent target [18], [19]. However, the available physical features for general SAR images are usually few. This hinders the application of more sophisticated warp function to model and correct more severe geometrical distortion. Therefore, the image-based features which can be sufficiently and automatically extracted from image although with less physical significance were proposed, such as the tie points [20]-[23]. Generally, any point pair that can geometrically tie two images can be the tie points. In SAR image registration, the tie points usually refer in particular to the features extracted from tie patches or regions. The tie patches are first matched by the region-based algorithms such as the cross-correlation (CC) algorithm [4], and then the tie points are constructed by extracting certain point positions from the matched patches such as the geometrical centers or centroids. Corner is another kind of image-based feature which is extracted by analyzing the local texture. A corner is a point which has two dominant and different edge directions in its local neighbourhood. Many corner detectors have been devised and Harris corner [24] is the commonly-used point feature for SAR image registration [7], [8], [21]. The Harris measure is the second moment matrix which describes the local neighbouring gradient distribution of a point. The corner response function of Harris is the weighted sum of the determinant and squared trace of the second moment matrix. A pixel is selected as a corner if its response is beyond a given threshold. The keypoint also names blob feature which refers to the point in image that

differs in properties like brightness or color compared with its surrounding [25]. Keypoint is extracted to further provide a complementary description of image structure in terms of region which cannot be obtained from corners [25]. The scale invariant feature transform (SIFT) [26] and the speeded up robust feature (SURF) [27] are the commonly-used keypoint features for SAR image registration. SIFT was proposed by Low to extract the scale-invariant feature based on Lindeberg's automatic scale selection theory [28]. Lindeberg found that the only possible scale space kernel is the Gaussian function under a variety of reasonable assumptions. Therefore, he experimented with both the trace of Hessian matrix, i.e. the Laplacian of Gaussian (LoG) and the determinant of Hessian (DoH) matrix to detect the blob-like structures. In order to efficiently extract the keypoints, Low proposed to simplify LoG with the difference of Gaussian (DoG) and showed that this approximation is successful because it improves the speed with just slight accuracy loss. SIFT is not only a feature detector but also a feature descriptor. The SIFT descriptor is a 128D vector of gradient and orientation. In [29], the authors presented a comparative study on ten different local descriptors and identified that SIFT is the most robust algorithm for treating common image deformations and performs the best. SIFT has been widely used for SAR image registration [30]-[35]. Chen et al. [32] systematically assessed the application of SIFT to SAR and showed its usefulness for image registration. Based on multisensor, multitemporal, and different viewpoint SAR images, Schwind et al. [34] further evaluated SIFT for SAR image registration and showed it is a potential robust alternative for point feature-based registration since subpixel consistency is achieved for most of the tested dataset. The bottleneck of SIFT is the speed [27], [32], [34], which hinders its application for general SAR image registration because the available SAR data is always large in size. In order to accelerate SIFT, Schwind et al. [34] proposed to skip features detected at the first octave of the scale space pyramid (SSP) because they found that there are only very few matches found at the first octave and the matches extracted from this octave have the highest matching false alarm rate (MFAR) among the all octaves for the selected SAR dataset. They showed that this can significantly reduce the detected keypoint number and processing time without reducing the correct match number. The first scale octave in the SSP of SIFT refers to the original or doubled image which has the highest resolution in the constructed SSP. The features extracted from this octave is more accurate for image registration because the localization precision at the higher scale octave is lower than that at the lower scale octave due to more smoothing thus coarser resolution [35]. Therefore, the discarding of matches from the first octave may impact the final registration accuracy. SURF proposed by Bay et al. [27] bases on the same scheme as SIFT but uses a combination of novel detection, description, and matching approaches to simplify the original algorithms to the essential. SURF extracts the feature based on DoH instead of its trace because DoH has slightly better scale selection property under non-Euclidean affine transformation than LoG [28], thus it fires less on elongated and ill-localized structures [27]. Bay et al. proposed a Fast-Hessian detector to approximate the DoH with box filters. The SURF descriptor is a 64D vector composed by the Harr wavelet responses of the square region around the keypoint and is also invariant to affine changes in illumination. For SAR image, this means the descriptor is invariant to the linear alteration of signal intensity. Therefore, the descriptor is robust to intensity undulations from imaging geometry, depolarization, and decorrelation in some sense. This indicates SURF is potential for registration of images from multisensor, multisource, and multimodality. Although both SURF and SIFT descriptors focus on the spatial gradient distribution, SURF is less sensitive to noise because it integrates the gradient information within a subpatch while SIFT depends on the orientation of the individual gradients [27]. SURF has been demonstrated to approximately or even outperform SIFT with respect to speed, repeatability, distinctiveness, and robustness [27]. Recently, SURF has been used in the remote sensing area for multispectral satellite image registration [36], seabed recognition based on the texture analysis of sonar images [37], as well as for SAR image registration [38], [39].

Although lots of feature-based registration algorithms for SAR images have been hitherto proposed, it seems that many algorithms are introduced from the optical image registration area in a sense. This poses some open problems which have not been perfectly solved. This paper is dedicated to investigate the appropriate feature for SAR image registration. We give a detailed evaluation to the commonly-used features for SAR image registration such as tie points, Harris corner, SIFT, and SURF in terms of geometrical invariance of feature, extraction speed, localization accuracy, geometrical invariance of descriptor, matching speed, robustness to decorrelation, and flexibility to image speckling. SURF is shown to be the best one. Particularly we find that SURF is flexible to image speckling due to the potential relation between Fast-Hessian detector and refined Lee speckle filter. It is further observed that by applying SURF on the oversampled SAR image and keeping the related sampling step unchanged, one can greatly improve the subpixel registration accuracy, the correct correspondence number, and the MFAR. Thus SURF is more appropriate for SAR image registration.

The rest paper is arranged as follows. Section 2 first evaluates the commonly-used features to index the appropriate feature for general SAR image registration, the identified optimal feature is then experimentally evaluated in Section 3 for high accurate image registration, and Section 4 finally concludes the paper.

2. EVALUATION OF COMMONLY-USED FEATURES FOR SAR REGISTRATION

In stead of developing a novel feature, here we identify the appropriate feature from the existing widely-used image-based features, i.e. the tie points, Harris corner, SIFT, and SURF by a comprehensive evaluation in terms of several key impact factors on SAR image registration, such as the geometrical invariance of feature, the extraction speed, the localization accuracy, the geometrical invariance of descriptor, the matching speed, the robustness to decorrelation, and the flexibility to image speckling.

The Geometrical Invariance of Feature. The invariance of feature indicates to which degree of warping the same feature can be still stably and distinctively extracted from the transformed image by the detector. We know that CC is sensitive to image rotation and scaling, thus the CC-based tie points are only invariant to translation. The second moment matrix of Harris measure is sensitive to image scaling, thus the extracted Harris corner is only invariant to translation and rotation. Because SIFT and SURF were proposed to further achieve the scale-invariance, they can guarantee stable feature extraction with higher order distortion compared with tie points and Harris. Although SIFT and SURF features are not full affine-invariant as those extracted by Harris-Affine and Hessian-Affine [40], it has been claimed that the affine frames in Harris-Affine and Hessian-Affine are more sensitive to noise than those of the scale-invariant detectors, so in practice the affine features have lower repeatability than the scale-invariant features unless the affine distortion is greater than about a 40 degree tilt of a planar surface [26], [40]. Large viewpoint change may result in large decorrelation and even different image content because the scattering is sensitive to the imaging geometry, thus it may impact the potential application. For the general SAR application, the scale-invariant features such as SIFT and SURF are sufficient.

The Feature Extraction Speed. As we know that the extraction speed is much affected by the computational load of each detector. The tie points are extracted by exhaustively traversing all potential offsets in two image directions to calculate CC. When conducting the subpixel extraction, the images should be further oversampled, which can lead to quadratic increasing in dataset, thus the resulted computational load is very heavy. The Harris point is extracted by analyzing the determinant and trace of the second moment matrix in each pixel position. The calculation of the second moment matrix only relates to the first order image derivative which can be easily obtained by just adding operations, thus Harris detector can perform very fast. SIFT and SURF extract the scale-invariant feature by first constructing the SSP, which consists of several octaves and each octave further consists of a constant number of scale levels. A scale level in SIFT is obtained by smoothing the image with a Gaussian. The next scale level is obtained by further applying the same filter to the output of the previous filtered layer. In this way we can obtain a series of scale levels. The nearby two layers are then subtracted to compute the DoG which is an approximation to the LoG. For the next octave, the images are subsampled and then the same procedure is performed. The final keypoint is extracted by selecting the point with extreme value of DoG using the non-maximum suppression in a $3 \times 3 \times 3$ neighbourhood in the scale space. SIFT detector performs slower than Harris because it extracts the feature in 3D space not in 2D space. But under the condition of extracting equal number of subpixel features, SIFT detector is still faster than CC-based tie points because the later conduct exhaustive search. SURF extracts the feature based on DoH. Given a point $\mathbf{x} = (x, y)$ in an image $\mathbf{I}$, the Hessian matrix $\mathbf{H}(\mathbf{x}, \sigma)$ in $\mathbf{x}$ at scale σ is defined as:

$$\mathbf{H}(\mathbf{x},\sigma) = \begin{pmatrix} L_{xx}(\mathbf{x},\sigma) & L_{xy}(\mathbf{x},\sigma) \\ L_{xy}(\mathbf{x},\sigma) & L_{yy}(\mathbf{x},\sigma) \end{pmatrix} \tag{1}$$

where $L_{xx}(\mathbf{x}, \sigma)$, $L_{yy}(\mathbf{x}, \sigma)$, and $L_{xy}(\mathbf{x}, \sigma)$ are respectively the convolution of the Gaussian second order derivative in x-, y-, and xy-directions with $\mathbf{I}$ in position $\mathbf{x}$. Thus the scale function DoH can be obtained by:

$$DoH = L_{xx}L_{yy} - \left(L_{xy}\right)^2. \tag{2}$$

When applied in the practice, Gaussians should be discretized and cropped. The corresponding discretized and cropped L_{xx}, L_{xy}, and L_{yy} with the lowest scale of 1.2 are shown in the top half of Figure 1. Inspired by the successful simplification of LoG with DoG in SIFT, Bay et al. proposed a Fast-Hessian detector to approximate the DoH with box filters. The box filters D_{xx}, D_{xy}, and D_{yy} respectively corresponding to L_{xx}, L_{xy}, and L_{yy} with scale of 1.2 are shown in the bottom half of Figure 1. It has been shown that the performance of this approximation is comparable or even better than the original discretized and cropped Gaussians [27]. From Figure 1 one can see that the approximation makes the pixels in certain window have the same weight. This enables the convolution can be calculated at very low computational cost by using the integral image. Another advantage of box filter is that the SSP can be constructed by applying the same box

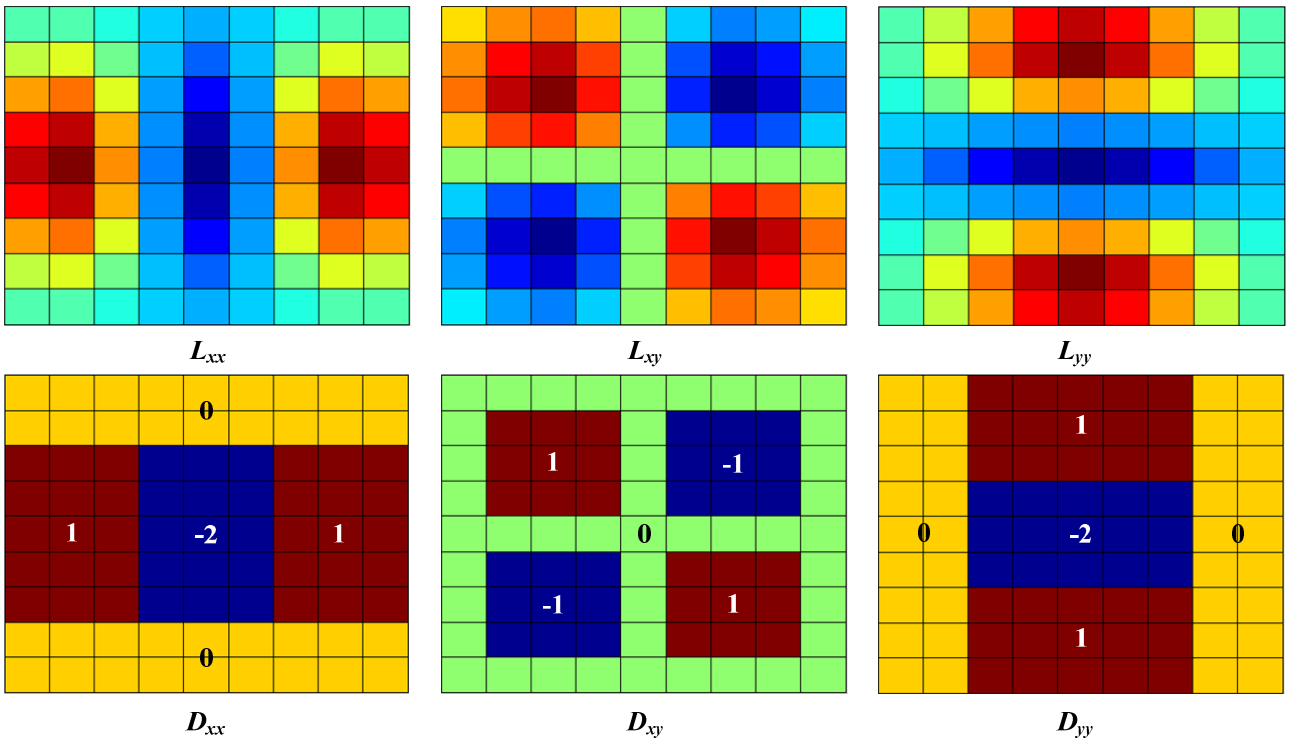

Figure 1. The discretized and cropped Gaussian second order partial derivatives respectively in *x*- (L_{xx}), *xy*- (L_{xy}), and *y*-direction (L_{yy}), as well as their corresponding SURF approximations D_{xx}, D_{xy}, and D_{yy} by box filter.

filters of any size directly on the original image at exactly the same speed because using of integral images enables the computation independent of the filter size. Therefore, instead of iteratively reducing the image size and using the cascade filtering, the SSP in SURF is built by up-scaling the filter size as shown in Figure 2. Thus SURF detector can perform much faster than SIFT. The scale function *DoH* in (2) can be then approximated as [27]

$$DoH_{approx} = D_{xx}D_{yy} - \left(0.9D_{xy}\right)^2. \tag{3}$$

Here the coefficient 0.9 is used to balance the expression for the Hessian's determinant. This function is then applied to locate the keypoint in the scale space using the same non-maximum suppression technique.

The Feature Localization Accuracy. The localization accuracy refers to how accurate the feature can be extracted, which is an important impact factor to the final registration accuracy. The tie points extracted by CC can achieve subpixel accuracy by oversampling the image patches [16], [41] or oversampling the CC map obtained in the coarse registration [4]. However, the obtained localization accuracy of tie points is dependent on the oversampling rate [23]. Higher sampling rate leads to higher accuracy but results in larger dataset thus heavier computational load. In fact, the oversampling rate should not be too high, because the increasing aliasing may then destroy the offset estimation. The accuracy of Harris corner is pixel, thus it can only be used for the coarse registration. The localization accuracies of SIFT and SURF are equally subpixel because they use the same localization approach. We know that SIFT and SURF extract the extrema by the non-maximum suppression technique, and the extrema is further located to subpixel and sub-scale accuracy by fitting a 3D quadratic to the scale function in the scale space. The scale functions are respectively the DoG and the approximated DoH for SIFT and SURF. The scale function f at a point $\mathbf{X} = (x, y, \sigma)$ in the scale space can be expanded in Taylor series up to the quadratic terms with $\mathbf{\Delta X}$ shifting from a detected extrema $\mathbf{X}_0 = (x_0, y_0, \sigma_0)$ [42]

$$f(\mathbf{X}) = f(\mathbf{X}_0) + \left(\frac{\partial f}{\partial \mathbf{X}}(\mathbf{X}_0)\right)^{\mathbf{T}} \mathbf{\Delta X} + \frac{1}{2}\mathbf{\Delta X}^{\mathbf{T}}\left(\frac{\partial^2 f}{\partial \mathbf{x}^2}(\mathbf{X}_0)\right)\mathbf{\Delta X} \tag{4}$$

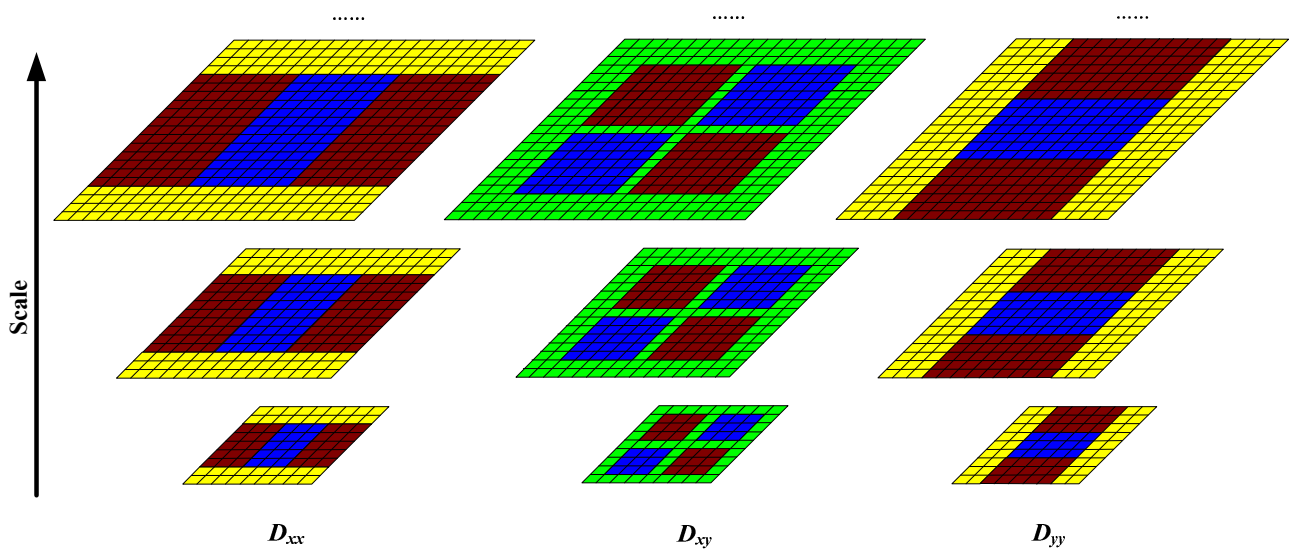

Figure 2. The scale space in SURF constructed by up-scaling the filter size.

where $\mathbf{\Delta X} = (\mathbf{X} - \mathbf{X}_0)$ is the offset from the extrema. The subpixel localization $\hat{\mathbf{\Delta X}}$ of the extrema is found by taking the derivative of (4) with respect to $\mathbf{\Delta X}$ and setting it to zero [42], [43]

$$\hat{\mathbf{\Delta X}} = -\left(\frac{\partial^2 f}{\partial \mathbf{X}^2}(\mathbf{X}_0)\right)^{-1}\left(\frac{\partial f}{\partial \mathbf{X}}(\mathbf{X}_0)\right) = -\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\mathbf{X}_0) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{X}_0) & \frac{\partial^2 f}{\partial x \partial \sigma}(\mathbf{X}_0) \\ \frac{\partial^2 f}{\partial x \partial y}(\mathbf{X}_0) & \frac{\partial^2 f}{\partial y^2}(\mathbf{X}_0) & \frac{\partial^2 f}{\partial y \partial \sigma}(\mathbf{X}_0) \\ \frac{\partial^2 f}{\partial x \partial \sigma}(\mathbf{X}_0) & \frac{\partial^2 f}{\partial y \partial \sigma}(\mathbf{X}_0) & \frac{\partial^2 f}{\partial \sigma^2}(\mathbf{X}_0) \end{pmatrix}^{-1}\begin{pmatrix} \frac{\partial f}{\partial x}(\mathbf{X}_0) \\ \frac{\partial f}{\partial y}(\mathbf{X}_0) \\ \frac{\partial f}{\partial \sigma}(\mathbf{X}_0) \end{pmatrix}. \quad (5)$$

The obtained $\hat{\mathbf{\Delta X}}$ will be accepted if it is less than 0.5 in all dimensions, so the final estimation of the extrema location can be obtained by adding $\hat{\mathbf{\Delta X}}$ to the location of its sampling position. Theoretically, the localization expression in (5) can be infinitely accurate without any limitation. It is to say SIFT and SURF can achieve the highest accuracy. However, it should be noted that although the subpixel feature localization is the precondition of accurate image registration, it cannot guarantee a subpixel image registration. Therefore, for high accurate SAR image registration, we should further evaluate the features, and this will be experimentally presented in Section 3.

The Geometrical Invariance of Descriptor. The feature descriptor is usually a vector to describe the neighbouring information of the feature. In image registration, feature descriptor is always used to construct the correspondences from the extracted features. The invariance of feature descriptor indicates to which degree of image warping two features can be still correctly matched. The tie points and Harris corner have no assigned descriptor. However, from the viewpoint of feature matching, both of them use the template matching, which means the image patch centered at the feature is used as the descriptor. The commonly-used image template is the squared window, which is only invariant to image translation. Thus the tie points and Harris corner can be only correctly matched in case of slight image distortion. SIFT and SURF descriptors can offer a good compromise between feature complexity and robustness to commonly occurring deformations, the affine transformation such as shearing and anisotropic scaling are also coved to some degree by the overall robustness of the descriptor [26], [27]. Thus the robust matching across a substantial range of affine distortion and change in 3D viewpoint can be achieved.

The Feature Matching Speed. The matching speed reflects the efficiency of correspondence construction. The feature matching is usually performed on certain merit function. A feature point in the master image is compared with a feature point in the slave image by calculating the merit function between their descriptors. For feature-based registration, the

merit function is usually selected as CC to maximize the similarity [4] or the Euclidean distance to minimize the differences [26], [27]. Generally, a correspondence is detected if it optimizes the merit function. For SIFT and SURF, an optimal correspondence is selected if its merit is also larger than certain times of the second optimal merit. Given the extracted features, the dominant factor affecting the matching speed is the calculation of the merit. For tie points and Harris corner, the merit function usually refers to CC, which can be computed with complex data or just magnitude data [41], referring to coherent CC or incoherent CC, respectively. It has been shown that the final registration accuracy achieved by using coherent CC is much higher than that by using incoherent CC [44]. If D_1 and D_2 are the image patches respectively centered at an initial match, the coherent CC is calculated as

$$CC(D_1, D_2) = \frac{\left| \sum_{i=1}^{N} \sum_{j=1}^{N} \left(D_1(i,j) - \mu_1 \right) \left(D_2(i,j) - \mu_2 \right)^* \right|}{\sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \left| \left(D_1(i,j) - \mu_1 \right) \right|^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \left| \left(D_2(i,j) - \mu_2 \right) \right|^2}} \tag{6}$$

where N is the size of the patch, μ_1 and μ_1 denote the means of D_1 and D_2, respectively. The superscript * denotes the complex conjugation. The zero-mean operation is necessary because it guarantees the obtained CC invariant to affine radiometric warping. The computation of (6) requires about $10N^2$ operations including $7N^2$ additions and $3N^2$ multiplications. For SIFT and SURF, the merit function is often the Euclidean distance. If D_3 and D_4 are respectively the descriptors of an initial match, the merit can be calculated by

$$Dist(D_3, D_4) = \sum_{i=1}^{L} \left| D_3(i) - D_4(i) \right|^2 \tag{7}$$

where L is the length of the descriptor. The computation of (7) requires about $3L$ operations including $2L$ additions and L multiplications. For SURF, Bay et al. [27] found that the sign of Laplacian can be further used to distinguish the feature from its background thus can be used for fast indexing during the matching stage. The merit will not be computed unless the considered initial match has the same sign of Laplacian. The two features which compose an initial match can be considered to be statistically independent because they are respectively extracted from different images. Therefore, under the assumption of the equal probability distribution of the positive and negative signs of Laplacian, the obtained merit computation for SURF will require $1.5L$ operations. Taking the descriptor lengths L for SIFT and SURF to be 128 and 64, respectively, thus the merit computation for SIFT and SURF will have respectively 384 operations and 96 operations, which means SURF is four times faster than SIFT on matching. In order to achieve the same computation as SIFT or SURF, the equivalent patch size N for tie points and Harris corner should be about 6 or 3, respectively. The small patch size may result in biased CC estimation thus bad feature localization and matching because of insufficient sampling.

The Robustness to Decorrelation. The decorrelation sources of SAR images can be classified into two categories, i.e. the geometrical warping and radiometric warping. The geometrical warping induces the decorrelation because CC is only invariant to translation. Thus the obtained correlation will be inaccurate and result in decorrelation for higher order geometrical warping. This can be related to the geometrical invariance of the feature, which has been discussed above. Here we only consider the radiometric warping-induced decorrelation. The radiometric warping introduces the decorrelation because CC is only invariant to affine change in scattering. In the microwave band, the target scattering is sensitive to frequencies, bandwidth, and polarization. Conversely, the scattering will also be different when there are some changes in the distributed targets such as the natural motions of water surface, vegetated lands, and snow-covered grounds, not to mention the obvious changes such as the developing foliage, the moving vehicle, and the constructing building. All of these introduce a complex nonlinear radiometric warping which decorrelates and degrades the SAR information thus aggravates the image registration difficulty. The image decorrelation badly impacts the localization accuracy of tie points because they are extracted based on correlation estimation. The achievable localization accuracy for tie points is given by the error standard deviation σ_L [17], [44], [45]:

$$\sigma_L = \sqrt{\frac{3}{2N^2}} \frac{\sqrt{1-\gamma^2}}{\pi\gamma} osr^{3/2} \tag{8}$$

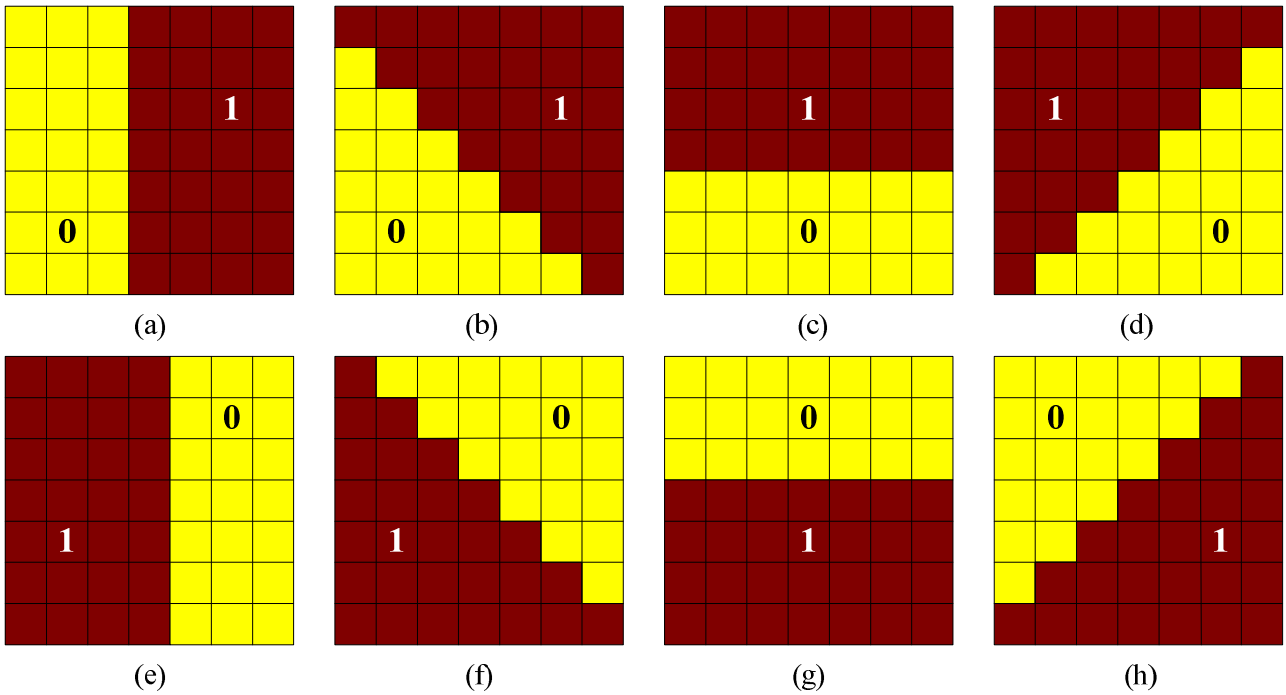

Figure 3. The eight edge-aligned windows used in refined Lee filter to decide the local texture, where the pixels in dark red are used in the filtering computation [46].

where γ is the CC of the two tie patches, N is the size of the patch, and osr is the oversampling rate of the data. Equation (8) indicates that the localization accuracy is directly related to the correlation. The higher coherence means higher localization accuracy, and the higher decorrelation means worse localization accuracy thus worse registration accuracy. We know that one can approximate the nonlinear function with a series of linear function and the zero-mean CC in (6) is invariant to affine radiometric warping, thus a good method to improve the robustness to decorrelation is to use smaller image patches, but this will result in worse localization accuracy in turn as shown in (8). Therefore, tie points are not robust to decorrelation. Similarly, the influence of decorrelation for CC-based matching of Harris corner is also unavoidable. However, Harris, SIFT, and SURF locate the feature based on geometrical texture such as the first and second order image derivatives in stead of correlation. In addition, the feature extractions in Harris as well as the lower scale octaves of SIFT and SURF only relate to the small neighbourhood around the feature, all of these will reduce the influence of decorrelation. For higher scale octaves of SIFT and SURF, the images are greatly smoothed thus the decorrelation will be reduced to some extent. The matching of SIFT and SURF feature is based on the local descriptors which are related to the local gradient and invariant to affine changes in scattering. Therefore, SIFT and SURF features are more robust to decorrelation.

The Flexibility to Image Speckling. SAR obtains the image by actively measuring and coherently processing the electromagnetic scattering signals from target. The coherent interference of scatterings reflected from different diffused scatterers contained within each resolution cell causes a pixel-to-pixel variation in intensities and results in the so-called speckle. For tie points based on image correlation, the assumption that the scattering is locally stationary and ergodic may not be completely tenable under the existence of speckle. Thus the correlation estimation may be biased which can result in inaccurate feature localization and incorrect feature matching. For the geometrical texture-based detector such as Harris, SIFT, and SURF, the speckle may lead to false texture thus high MFAR. In order to extract stable feature from the speckle-contaminated SAR image, a conceivable approach is to suppress the speckle with filter before extracting features. Schwind et al. [34] proposed to use the ISEF filter to reduce the influence of speckle, however, they also pointed out that image smoothing by an ISEF filter or any other speckle filter may slightly affect feature localization and finally the registration quality, so it may be very difficult to estimate without real ground truth. Therefore, a perfect

Table 1. Evaluation of the four commonly-used features for SAR image registration in terms of several criteria.

Items	Tie Points	Harris Corner	SIFT	SURF
Geometrical Invariance of Feature	Translation	Rotation and translation	Scaling, rotation, and translation	Scaling, rotation, and translation
Feature Extraction Speed	Slower	Faster	Slow	Fast
Feature Localization Accuracy	Subpixel but controlled by the sampling rate	Pixel	Subpixel	Subpixel
Geometrical Invariance of Feature Descriptor	Translation	Translation	Affine transformation	Affine transformation
Feature Matching Speed	Slow	Slow	Fast	Faster
Robustness to Decorrelation	Worse	Bad	Good	Good
Flexibility to Image Speckling	Good	Bad	Bad	Better

strategy is to conduct speckle filtering along with feature extraction. As for feature extraction, the detector is thus required to be flexible to image speckling.

The Harris detector extracts the features directly based on the first order image derivative, so it is hard to be immune from image speckling. Therefore, Harris corner may extract many features from the images but has only a small number of correct matches, because most of the extracted features are speckles. This influence is also observed for SIFT by Schwind et al. [34] when evaluating the applicability of SIFT for SAR. They found that there are only very few matches found at the first octave of SSP although with extensive number of extractable features, and the matches extracted from this octave have the highest MFAR of the all octaves. For the higher octaves, this influence can be weakened because the lager image smoothing may reduce the speckle to some extent. The first scale octave refers to the original or double-sized images which are of the highest resolution and the largest number extractable feature points. The few matches and the highest MFAR at this octave clearly indicates the bad flexibility of SIFT to speckle. However, SURF can deal with speckle very well because the potential relation between Fast-Hessian detector and the refined Lee speckle filter, as will be shown in the following.

The primary goal of speckle filtering is to reduce the speckle without sacrificing the image content. The most commonly-applied technique to reduce speckle is the boxcar filter, which replaces a pixel with the average of its windowed neighbourhood. This filter can be easily implemented and works very well for homogeneous areas. However, it has a major drawback of spatial resolution degradation due to indiscriminately averaging pixels from inhomogeneous media [54], thus the edges may be blurred and the bright point targets may be smeared. The ideal speckle filter should adaptively smooth the speckle, retain the sharpness of edge and feature boundaries, as well as preserve the subtle but distinguishable details such as thin linear features and point targets [46]. To achieve this goal, many other filtering techniques have been proposed and among which the refined Lee filter is one of the most commonly-used. Lee filter [47] uses the local statistics such as the mean and variance to conduct the despeckling. In order to reduce speckle without degrading the image, the neighbouring pixels with the similar texture characteristics as the center pixel are selected. Lee proposed to match the edge direction using eight nonsquare edge-aligned windows as shown in Figure 3. In the course of filtering, one of the nonsquare windows is selected based on the edge direction to calculate the local statistics, and then the minimum mean square algorithm is applied. It has been shown that this filter can effectively reduce the speckle without degrading the edge [46], [47].

As mentioned previously, SURF extracts the feature based on DoH, which can be simplified by the box filter. The using of box filter not only accelerates the feature extraction but also provides a perfect approach to extract the edge features while reducing the speckle. The D_{xx} in Figure 1 indicates that we can reduce the speckle by using a 5 × 3 window to average the pixels and then extract the vertical edge feature by the second order image partial derivative in x-direction with the convolution template of [1 -2 1]. This is equivalent to filter the speckle with Lee's nonsquare windows (a) and (e). Similarly, the D_{yy} indicates that we can reduce the speckle by using a 5 × 3 window and extract the horizontal edge feature by the second order image partial derivative in y-direction with the convolution template of $[1\ -2\ 1]^{T}$. This is equivalent to filter the speckle with Lee's nonsquare windows (c) and (g). The D_{xy} indicates that we can reduce the speckle by using a 3 × 3 window and extract the 135° edge feature by the second order image partial derivative in

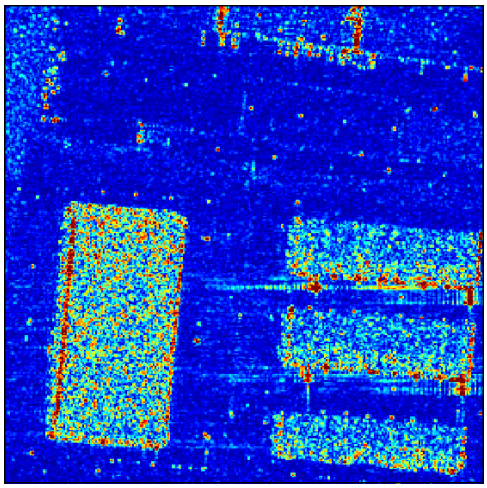

Figure 4. The SAR image (300×300) taken by RadarSat-2.

negative *xy*-direction with the convolution template of $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. This is equivalent to filter the speckle with Lee's nonsquare windows (d) and (h). The $-D_{xy}$ indicates that we can reduce the speckle by using a 3 × 3 window and extract the 45° edge feature by the second order image partial derivative in *xy*-direction with the convolution template of $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$. This is equivalent to filter the speckle with Lee's nonsquare windows (b) and (f). In stead of selecting the optimal edge to estimate the local statistical, the all four speckle-reduced edge features in SURF are combined to a new characteristic value through $D_{xx} \times D_{yy} + 0.9D_{xy} \times 0.9(-D_{xy})$, which corresponds to the approximated DoH in (3) and will be used for further keypoint localization. The SSP shown in Figure 2 indicates that we can use a series of windows of different size to filter the speckle and extract the features of different scales. Therefore, SURF is very flexible to deal with speckle.

According to the above analysis, we can give a comprehensive evaluation on the commonly-used point features for SAR image registration in terms of several criteria and the evaluation results are listed in Table 1, from which we can summarize for general SAR image registration that:

- SURF behaves the best in terms of the selected evaluation criteria.
- SIFT may be applicable when no strict requirement for speed.
- Harris is only good for coarse registration.
- Tie points are appropriate for slightly distorted and decorrelated images but need heavy computation.

3. EVALUATION OF SURF FOR SAR IMAGE SUBPIXEL REGISTRATION

From the above evaluation we can draw the conclusion that the SURF feature is more appropriate and competent for general SAR image registration. Nevertheless, we need to further evaluate SURF with real data, especially on the registration accuracy. We know that the SAR applications usually have a strict requirement for registration accuracy, especially for interferometric SAR-based elevation or deformation estimation. To ensure an acceptable estimation, the registration should achieve subpixel accuracy [48]. In order to evaluate the capability of SURF for subpixel image registration, we devise a comparative experiment on some contrived SAR image pairs. Figure 4 shows a SAR image taken from South Phoenix, AZ by RadarSat-2 on May 4, 2008. We treat this image as the master and the one obtained by affine warp transforming as the slave image through:

$$\begin{pmatrix} x_s \\ y_s \\ 1 \end{pmatrix} = \mathbf{A}\begin{pmatrix} x_m \\ y_m \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_m \\ y_m \\ 1 \end{pmatrix} \tag{9}$$

where $(x, y, 1)^{\mathrm{T}}$ denotes the homogenous image coordinates, the subscript s and m indicate the slave and master, respectively. $\mathbf{A}$ is an affine matrix with parameters a, b, c, d, as well as two translations t_x and t_y. Thus we obtain an image pair with controllable registration parameters. Bay et al. proposed two different versions of Fast-Hessian detector for SURF, the one with an initial filter size of 9 × 9 on the original image is denoted as *FH*-9(-1), and the one with an initial filter size of 15 × 15 on the doubled image with doubled sampling step is denoted as *FH*-15(-2). The number in the parentheses indicates the image size: "1" denotes the extraction on the original image and "2" denotes the extraction on the doubled image. It has been shown that *FH*-15(-2) has a better performance than *FH*-9(-1) in terms of repeatability [27]. We respectively apply the two versions of detector to extract the point correspondences, based on which the robust extended fast least trimmed squares [49] is used to retrieve the warp matrix. In order to compare the two detectors for SAR registration, four criteria, i.e. the average transfer error (*ATE*), the number of correct matches, the *MFAR*, and the warp matrix estimation error (*WMEE*) are considered. *ATE* is used to evaluate the consistency of the extracted features with the retrieval parameters, and it is defined as:

$$ATE = \frac{1}{N}\sum_{i=1}^{N}\left\| \begin{pmatrix} x_{si} \\ y_{si} \\ 1 \end{pmatrix} - \hat{\mathbf{A}}\begin{pmatrix} x_{mi} \\ y_{mi} \\ 1 \end{pmatrix} \right\| \tag{10}$$

where $\hat{\mathbf{A}}$ denotes the estimated warp matrix on the all constructed correspondences, (x_{si}, y_{si}) and (x_{mi}, y_{mi}) denote the ith correct correspondence respectively located in the slave and master images, N is the number of correct matches which are selected by:

$$\left\| \begin{pmatrix} x_{si} \\ y_{si} \\ 1 \end{pmatrix} - \mathbf{A}\begin{pmatrix} x_{mi} \\ y_{mi} \\ 1 \end{pmatrix} \right\| < threshold \quad \begin{cases} \text{True} & (x_{si}, y_{si}) \leftrightarrow (x_{mi}, y_{mi}) \text{ is a correct match} \\ \text{False} & (x_{si}, y_{si}) \leftrightarrow (x_{mi}, y_{mi}) \text{ is a mismatch} \end{cases} \tag{11}$$

where $\mathbf{A}$ is the true warp matrix. The *threshold* in the experiment is chosen as 5 pixels, which indicates that a correspondence is treated as a mismatch if the transfer error of it in any image direction is larger than 5 pixels. The *MFAR*, also called 1-precision [29], is defined as:

$$MFAR = \frac{\#matches - \#correct\ matches}{\#matches} \tag{12}$$

where "#" means "the number of". From (12) one can see that *MFAR* refers to the percentage of mismatches among the constructed correspondences and it is mainly influenced by the geometrical warping, radiometric warping, and speckle. For a given contrived SAR image pair with controlled geometrical and radiometric warping, the combination of correct match number and *MFAR* can be used to evaluate the flexibility of the detector to image speckling. The *WMEE* as defined in (13) is used to further evaluate the consistency and accuracy of the estimated warp matrix with its true value:

$$WMEE = \left\| \mathbf{A} - \hat{\mathbf{A}} \right\|_F \tag{13}$$

where F indicates the Frobenius norm.

We exemplify four image pairs with different transformations, the estimated registration parameters as well as the *ATE*, correct match number, *MFAR*, and *WMEE* for the two detectors are listed in Table 2, from which one can see that the *FH*-15(-2) can extract more correct matches with lower *MFAR* than *FH*-9(-1). It means SURF is more flexible to image speckling, because *FH*-15(-2) performs the feature extraction on the doubled image with more serious speckle. The *ATE* of *FH*-15(-2) is smaller than that of *FH*-9(-1). This demonstrates that the *FH*-15(-2) features are more consistent with the retrieval parameters. As for the *WMEE*, the *FH*-15(-2) does not improve the registration accuracy for the all four cases as we expected. There is still clear inconsistency between the estimated warp matrix and its true value. We think the main

Table 2. Evaluation of each Fast-Hessian detectors on the four SAR image pairs in terms of the retrieval registration parameters, the number of correct matches, the matching false alarm rate (*MFAR*), the average transfer error (*ATE*), and the warp matrix estimation error (*WMEE*).

Detectors	Registration Parameters						# CorrectMatch and *MFAR*	*ATE*	*WMEE*
	a	*b*	*c*	*d*	t_x	t_y			
True Value	0.7189	0.0452	-0.0402	0.8087	1.7000	2.4000	--	--	--
FH-15(-2)	0.7193	0.0458	-0.0405	0.8103	1.2700	1.9548	181 (0.1381)	(0.5906, 0.5930)	0.6189
FH-9(-1)	0.7180	0.0462	-0.0382	0.8124	1.3689	1.1133	74 (0.2600)	(0.7019, 0.8956)	1.3286
FH-9(-2)	0.7192	0.0449	-0.0401	0.8087	1.5326	2.1988	353 (0.1108)	(0.3249, 0.2847)	0.2617
FH-9(-3)	0.7188	0.0452	-0.0402	0.8087	1.5414	2.2627	920 (0.0791)	(0.2210, 0.1820)	0.2098
FH-9(-4)	0.7189	0.0452	-0.0402	0.8087	1.5527	2.2729	1379 (0.0561)	(0.1754, 0.1561)	0.1946
FH-9(-5)	0.7189	0.0451	-0.0403	0.8088	1.5573	2.2607	1760 (0.0517)	(0.1546, 0.1438)	0.1994
True Value	0.9361	0.1889	-0.1617	1.0938	-10.5000	-3.4000	--	--	--
FH-15(-2)	0.9368	0.1895	-0.1622	1.0945	-10.6192	-3.4582	235 (0.0891)	(0.5704, 0.6250)	0.1327
FH-9(-1)	0.9352	0.1914	-0.1616	1.0926	-10.7810	-3.1672	99 (0.1818)	(0.7493, 0.8038)	0.3649
FH-9(-2)	0.9361	0.1890	-0.1616	1.0938	-10.4161	-3.4555	568 (0.0486)	(0.3104, 0.2486)	0.1006
FH-9(-3)	0.9361	0.1888	-0.1617	1.0938	-10.3958	-3.4413	1579 (0.0419)	(0.1790, 0.1636)	0.1121
FH-9(-4)	0.9361	0.1889	-0.1617	1.0938	-10.4180	-3.4346	2090 (0.0355)	(0.1576, 0.1471)	0.0890
FH-9(-5)	0.9362	0.1889	-0.1617	1.0938	-10.4307	-3.4547	2440 (0.0360)	(0.1584, 0.1417)	0.0883
True Value	1.1365	0.1036	-0.0894	1.3159	-2.6000	5.4000	--	--	--
FH-15(-2)	1.1376	0.1023	-0.0898	1.3183	-2.3666	5.3449	169 (0.1675)	(0.6171, 0.6243)	0.2398
FH-9(-1)	1.1355	0.1030	-0.0897	1.3200	-2.2357	5.1394	70 (0.2135)	(0.7558, 0.8803)	0.4479
FH-9(-2)	1.1366	0.1037	-0.0894	1.3164	-2.4500	5.5012	378 (0.0892)	(0.3013, 0.3085)	0.1810
FH-9(-3)	1.1365	0.1036	-0.0895	1.3160	-2.4446	5.5492	1120 (0.0517)	(0.1765, 0.1768)	0.2154
FH-9(-4)	1.1365	0.1036	-0.0894	1.3159	-2.4521	5.5446	1627 (0.0568)	(0.1373, 0.1313)	0.2068
FH-9(-5)	1.1365	0.1035	-0.0895	1.3159	-2.4530	5.5459	1896 (0.0472)	(0.1335, 0.1176)	0.2071
True Value	1.2079	0.0777	-0.0718	1.3077	-5.3000	1.5000	--	--	--
FH-15(-2)	1.2073	0.0768	-0.0716	1.3088	-4.8525	1.6586	170 (0.1327)	(0.5590, 0.5550)	0.4748
FH-9(-1)	1.2074	0.0769	-0.0705	1.3096	-4.9986	1.1952	73 (0.2065)	(0.8510, 0.7991)	0.4287
FH-9(-2)	1.2081	0.0775	-0.0715	1.3079	-5.0778	1.6065	396 (0.0682)	(0.2693, 0.2545)	0.2464
FH-9(-3)	1.2079	0.0777	-0.0719	1.3077	-5.1180	1.6681	1155 (0.0478)	(0.1710, 0.1794)	0.2477
FH-9(-4)	1.2079	0.0777	-0.0718	1.3077	-5.1227	1.6514	1702 (0.0518)	(0.1225, 0.1139)	0.2331
FH-9(-5)	1.2079	0.0777	-0.0717	1.3077	-5.1326	1.6386	1963 (0.0549)	(0.1206, 0.1051)	0.2173

The value in the parentheses of the eighth column refers to the *MFAR*. The values in the parentheses of the ninth column refer to the *ATE* in *x*- and *y*-directions, respectively.

reason is when *FH*-15(-2) doubles the image size the sampling step is also doubled. Although with different integral image, the feature detection of *FH*-15(-2) can be geometrically equivalent to that on the original image with the initial filter size of 7.5 × 7.5, i.e. *FH*-7.5(-1). Smaller filter size indicates higher resolution and more extractable features, as shown in Table 2. But the doubled sampling step makes the sampling is still conducted on the same and equivalent pixel position not on the subpixel image position. For example, if (x_0, y_0) is a sampled pixel position in the original image, the corresponding position of this pixel in the doubled image is $(2x_0, 2y_0)$, the doubled sampling step makes this pixel position is still sampled but not on the $(2x_0 \pm 1, 2y_0 \pm 1)$, which correspond the subpixel position $(x_0 \pm 0.5, y_0 \pm 0.5)$ in the original image. These subpixel features are of doubled position accuracy and contribute more to the subpixel image registration. Therefore, we propose to conduct the detection with the initial filter size of 9 × 9 and unchanged sampling step but on the oversampled images, we denote this detector as *FH*-9(-*Fs*), where *Fs* indicates the sampling rate. We recommend using the linear interpolator such as the bilinear interpolator to conduct the sampling to avoid the nonlinear image aliasing. Table 2 further lists the registration results based on the *FH*-9(-2) to *FH*-9(-5) detectors. We can see that the *ATE*, correct match number, *MFAR*, and *WMEE* of *FH*-9(-2) are perfectly improved comparing with that of *FH*-9(-1) and *FH*-15(-2), at the same time the registration accuracy is also improved as the oversampling rate increases because more feature correspondences can be extracted with higher localization accuracy. This makes the SAR image registration with higher accuracy possible. The image oversampling will increase the dataset thus the computational load but it is still acceptable. For high accuracy registration, we recommend to oversample the image three or four times so as to achieve the compromise among accuracy, robustness, and computational complexity.

4. CONCLUSION

The imaging geometry and mechanism of SAR inherently bring about the unavoidable geometrical distortion and speckle to the obtained images, which makes SAR image registration much complicated compared with that of optical images. It seems that many existing feature-based algorithms are introduced from the optical image registration area in a sense. This poses some open problems which have not been perfectly solved. This paper is dedicated to investigate the appropriate feature for SAR image registration. We give a comprehensive evaluation to the commonly-used features such as tie points, Harris corner, SIFT, and SURF in terms of several criteria such as the geometrical invariance of feature, the extraction speed, the localization accuracy, the geometrical invariance of descriptor, the matching speed, the robustness to decorrelation, and the flexibility to image speckling. It is shown that SURF outperforms others. Among these criteria, we particularly address the feature's flexibility to the unavoidable SAR speckle which degrades the image information and should be suppressed. Since the speckle filtering may change the feature position and impact the subpixel feature localization in turn, thus a good feature detector should be robust to speckle. We find that the Fast-Hessian detector of SURF has a potential relation with the refined Lee speckle filter, which indicates that SURF can extract image feature at different scales even under the influence of image speckling. This observation is validated when we oversample the image with increasing sampling rate, i.e. refine the image pixel, it means much more severe speckle would be suffered, unlike SIFT, the numbers of extractable correct correspondences by SURF and the matching false alarm rate, however, are still improved accordingly. In consideration of some applications with strict requirement for registration accuracy, we suggest using the original Fast-Hessian detector on the oversampled image with unaltered sampling step to extract the feature. The registration experiments on some SAR image pairs demonstrate that the suggested SURF can significantly improve the subpixel registration accuracy and speckle immunity. Therefore, we think SURF is more appropriate and competent for general SAR image registration.

REFERENCES

[1] Li, H., Manjunath, B. S. and Mitra, S. K., "A contour-based approach to multisensor image registration," IEEE Trans. Image Process. 4(3), 320-334 (1995).

[2] Meddeber, L., Berrached, N. E. and Taleb-Ahmed, A., "An automatic registration and mosaicking system based on contour features and wavelet transformation for remote sensing images," Proc. Int. Conf. Signals, Circuits Syst., 1-7 (2009).

[3] Gabriel, A. K. and Goldstein, R. M., "Crossed orbit interferometry: Theory and experimental results from SIR-B," Int. J. Remote Sens. 9(5), 857-872 (1988).

[4] Li, F. K. and Goldstein, R. M., "Studies of multibaseline spaceborne interferometric synthetic aperture radars," IEEE Trans. Geosci. Remote Sens. 28(1), 88-97 (1990).

[5] Sakurai-Amano, T. *et al.*, "Scene feature matching analysis of JERS-1 SAR images," Proc. IGARSS 2, 1037-1039 (1998).

[6] Zalmanson, G. H., Doytsher, Y. and Schenk, T., "Towards a new paradigm for registration of SAR images using linear features," Proc. ISPRS Congr. 3, 41-45 (2004).

[7] Marinkovic, P. and Hanssen, R. F., "Advanced InSAR coregistration using point clusters," Proc. IGARSS 1, 489-492 (2004).

[8] Bentoutou, Y., Taleb, N., Kpalma, K. and Ronsin, J., "An automatic image registration for applications in remote sensing," IEEE Trans. Geosci. Remote Sens. 43(9), 2127-2137 (2005).

[9] Dare, P. M. and Dowman, I. J., "Automatic registration of SAR and SPOT imagery based on multiple feature extraction and matching," Proc. IGARSS 7, 2896-2898 (2000).

[10] Dare, P. M. and Dowman, I. J., "An improved model for automatic feature-based registration of SAR and SPOT images," ISPRS J. Photogramm. Remote Sens. 56(1), 13-28 (2001).

[11] Growe, S. and Tonjes, R., "A knowledge based approach to automatic image registration," Proc. IEEE Int. Conf. Image Process. 3, 228–231 (1997).

[12] Dell'Acqua, F., Gamba, P. and Lisini, G., "Coregistration of multiangle fine spatial resolution SAR images," IEEE Geosci. Remote Sens. Lett. 1(4), 237-241 (2004).

[13] Zhang, X., Xi, R., Li, H., Zhang, X. and Yang, T., "Image registration based on rectangle pattern," Proc. ICIG, 319-324 (2010).
[14] Gelautz, M., Mitteregger, E. and Leberl, F., "Automated acquisition of ground control using SAR layover and shadows," Proc. IGARSS 1, 468-470 (1997).
[15] Toutin, T., Hoja, D., Hoeppner, E., Remond, A. and King, C., "GCPs selection from multi-source data over mountainous topography," Proc. IGARSS 5, 2339-2341 (1998).
[16] Rufino, G., Moccia, A. and Esposito, S., "DEM generation by means of ERS tandem data," IEEE Trans. Geosci. Remote Sens. 36(6), 1905–1912 (1998).
[17] Serafino, F., "SAR image coregistration based on isolated point scatterers," IEEE Geosci. Remote Sens. Lett. 3(3), 354-358 (2006).
[18] Zhang, L., Ding, X. and Lu, Z., "Ground settlement monitoring based on temporarily coherent points between two SAR acquisitions," ISPRS J. Photogramm. Remote Sens. 66(1), 146-152 (2011).
[19] Perissin, D. and Wang, T., "Repeat-pass SAR interferometry with partially coherent targets," IEEE Trans. Geosci. Remote Sens. 50(1) 271–280 (2012).
[20] Homer, J. and Longstaff, I. D., "Minimising the tie patch window size for SAR image coregistration," Electron. Lett. 39(1), 122-124 (2003).
[21] Zou, W., Li, Y., Li, Z. and Ding, X., "Improvement of the accuracy of InSAR image coregistration based on tie points: A review," Sensors 9(2), 1259-1281 (2009).
[22] Giles, A. B., Massom, R. A. and Warner, R. C., "A method for sub-pixel scale feature-tracking using Radarsat images applied to the Mertz Glacier Tongue, East Antarctica," Remote Sens. Environ. 113(8), 1691-1699 (2009).
[23] Li, D. and Zhang, Y., "A fast offset estimation approach for InSAR image subpixel registration," IEEE Geosci. Remote Sens. Lett. 9(2), 267-271 (2012).
[24] Harris, C. and Stephens, M., "A combined corner and edge detector," Proc. Alvey Vis. Conf., 141-151 (1988).
[25] "Blob Detection," < http://en.wikipedia.org/wiki/Blob_detection>.
[26] Lowe, D. G., "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis. 60(2), 91-110 (2004).
[27] Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., "Speeded-up robust features (SURF)," Comput. Vis. Image Understand. 110(3), 346–359 (2008).
[28] Lindeberg, T., "Feature detection with automatic scale selection," Int. J. Comput. Vis. 30(2), 77-116 (1998).
[29] Mikolajczyk, K. and Schmid, C., "A performance evaluation of local descriptors," IEEE Trans. Pattern Anal. Mach. Intell. 27(10), 1615-1630 (2005).
[30] Wessel, B., Huber, M. and Roth, A., "Automatic image-to-image registration of near real-time SAR images," Proc. ENVISAT Symp., SP-636 (2007).
[31] Li, F., Zhang, G. and Yan, J., "Coregistration based on SIFT algorithm for synthetic aperture radar interferometry," Proc. ISPRS Congr. 1, 123-128 (2008).
[32] Chen, E. -X., Li, Z. -Y., Tian, X. and Li, S. -M., "Application of scale invariant feature transformation to SAR imagery registration," Acta Automatica Sinica 34(8), 861-868 (2008).
[33] Li, D. and Zhang, Y., "Geometric feature-based image coregistration approach for InSAR," Proc. APSAR, 1026-1030 (2009).
[34] Schwind, P., Suri, S., Reinartz, P. and Siebert, A., "Applicability of the SIFT operator to geometric SAR image registration," Int. J. Remote Sens. 31(8), 1959-1980 (2010).
[35] Wang, S., You, H. and Fu, K., "BFSIFT: A novel method to find feature matches for SAR image registration," IEEE Geosci. Remote Sens. Lett. 9(4), 649-653 (2012).
[36] Teke, M. and Temizel, A., "Multi-spectral satellite image registration using scale-restricted SURF," Proc. Int. Conf. Pattern Recognit., 2310-2313 (2010).
[37] Nguyen, H. -G., Fablet, R., Ehrhold, A. and Boucher, J. -M., "Keypoint-based analysis of sonar images: Applications to seabed recognition," IEEE Trans. Geosci. Remote Sens. 50(4), 1171-1184 (2012).
[38] Liu, R. and Wang, Y., "SAR image matching based on speeded up robust feature," Proc. WRI Global Congr. Intell. Syst. 4, 518-522 (2009).
[39] Lee, W. -K. and Kim, A. -L., "An efficient automatic geo-registration technique for high resolution spaceborne SAR image fusion," Proc. IGARSS, 3566-3569 (2011).
[40] Mikolajczyk, K. and Schmid, C., "An affine invariant interest point detector," Proc. ECCV 1, 128-142 (2002).
[41] Li, Z. and Bethel, J., "Image coregistration in SAR interferometry," Proc. ISPRS Congr. 1, 433-438 (2008).

[42] Brown, M. and Lowe, D. G., "Invariant features from interest point groups," Proc. BMVC, 253-262 (2002).
[43] Zhang, N., "Computing parallel speeded-up robust features (P-SURF) via POSIX threads," Proc. Int. Conf. Intell. Comput., 287-296 (2009).
[44] Bamler, R. and Eineder, M., "Accuracy of differential shift estimation by correlation and split-bandwidth interferometry for wideband and delta-k SAR systems," IEEE Geosci. Remote Sens. Lett. 2(2), 151-155 (2005).
[45] Bamler, R., "Interferometric stereo radargrammetry: Absolute height determination from ERS-ENVISAT interferograms," Proc. IGARSS 2, 742-745 (2000).
[46] Lee, J. -S. and Pottier, E., [Polarimetric Radar Imaging: From Basics to Applications], CRC Press, Boca Raton, FL, 143-177 (2009).
[47] Lee, J. -S., "Refined filtering of image noise using local statistics," Comput. Graphics Image Process. 15(4), 380-389 (1981).
[48] Hanssen, R. F., [Radar Interferometry: Data Interpretation and Error Analysis], Kluwer Academic Publishers, Dordrecht, The Netherlands, 45-47 (2001).
[49] Li, D. and Zhang. Y., "The appropriate parameter retrieval algorithm for feature-based SAR image registration," SPIE Conf. SAR Image Anal. Model. Tech., 2012.