We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

On Feature-Based SAR Image Registration: Appropriate Feature and Retrieval Algorithm

Dong Li, Yunhua Zhang and Xiaojin Shi

Abstract

An investigation on the appropriate feature and parameter retrieval algorithm is conducted for feature-based registration of synthetic aperture radar (SAR) images. The commonly used features such as tie points, Harris corner, SIFT, and SURF are comprehensively evaluated. SURF is shown to outperform others on criteria such as the geometrical invariance of feature and descriptor, the extraction and matching speed, the localization accuracy, as well as the robustness to decorrelation and speckling. The processing result reveals that SURF has nice flexibility to SAR speckles for the potential relationship between Fast-Hessian detector and refined Lee filter. Moreover, the use of Fast-Hessian to oversampled images with unaltered sampling step helps to improve the registration accuracy to subpixel (i.e., <1 pixel). As for parameter retrieval, the widely used random sample consensus (RANSAC) is inappropriate because it may trap into local occlusion and result in uncertain estimation. An extended fast least trimmed squares (EF-LTS) is proposed, which behaves stable and averagely better than RANSAC. Fitting SURF features with EF-LTS is hence suggested for SAR image registration. The nice performance of this scheme is validated on both InSAR and MiniSAR image pairs.

Keywords: extended fast least trimmed squares (EF-LTS), feature-based image registration, parameter estimation, speeded up robust feature (SURF), synthetic aperture radar (SAR)

1. Introduction

Synthetic aperture radar (SAR) as an irreplaceable remote sensing technique has been used for earth observation and environment monitoring for a long time due to its all-weather and all-day operational capability. A large number of airborne and spaceborne SAR sensors have been deployed recently. Nevertheless, the difference in sensors and imaging geometries will always introduce a geometrical warp between images which should be compensated before any joint application of multiple SAR images for accurate apperception and understanding of target and scene. Image registration is just dedicated to retrieve the warp function to align the same pixel position in each SAR image to the same target in the global system.

A lot of SAR image registration techniques have been developed hitherto. In this chapter, we focus on the algorithms that conduct registration based on image features, such as contour, region, line, and point. Contour, region, and line as well

as their combination are often used for registration of multi-modality images. For SAR images with geometrical distortion and speckle, point feature is generally much clearer and easier extracted. Tie points, corner, and keypoint are the commonly used features in SAR image registration. Tie points usually refer to the features extracted from tie patches in SAR image registration [1–4]. The tie patches are first matched by region-based algorithms, and the tie points are then located by extracting the geometrical centers or centroids of the matched patches. Corner denotes another kind of point feature which has two dominant but different edge directions in local neighborhood. In SAR image registration, Harris corner [5] is the commonly used point feature [2, 6] whose response function is the weighted addition of the determinant and squared trace of the first-order moment matrix which describes the local neighboring gradient distribution of a point. Keypoint refers to the point differing in brightness or color compared with the surrounding. It is identified to further enable a complementary description of image structure that cannot be characterized by corner. The scale invariant feature transform (SIFT) [7] and the speeded up robust feature (SURF) [8] are the widely used keypoints in SAR image registration. SIFT was developed by Lowe [7] to extract features based on the automatic scale selection theory. Lindeberg [9] found that the only possible scale-space kernel under a variety of reasonable assumptions is the Gaussian function, and he experimented with both the traces of Hessian matrix, i.e., the Laplacian of Gaussian (LoG) and the determinant of Hessian (DoH) matrix, to detect the blob-like structures. To extract keypoints efficiently, Lowe [7] simplified LoG with the difference of Gaussian (DoG) further. SIFT enables not only a feature detector, but also a 128D vectorized descriptor of gradient and orientation. Mikolajczyk and Schmid conducted a comparative study on 10 different local descriptors and found that SIFT performs the best on treating the common image deformations [10]. SIFT has been widely used in SAR image registration [11–23]. Chen et al. [13] systematically evaluated the application of SIFT to SAR and displayed its usefulness for image registration. Schwind et al. [15] further indicated that SIFT is a robust alternative for point feature-based SAR image registration. The bottleneck of SIFT is the speed [8, 13, 15], which hinders its application to general SAR image registration. To accelerate SIFT, Schwind et al. [15] proposed to skip features detected at the first octave of the scale space pyramid (SSP) because matches extracted from this octave have the highest matching false alarm rate (MFAR). This can save the processing time without reducing the number of correct matches greatly. However, the first scale octave in SSP of SIFT refers to the image of original size or doubled size which has the highest resolution in SSP. Thus, the features extracted from this octave are more accurate for image registration [16]. Therefore, the discarding of matches from the first octave may influence the final registration accuracy. Based on the same scheme as SIFT, SURF developed by Bay et al. [8] uses a combination of novel detection, description, and matching methods to simplify SIFT. SURF extracts feature based on DoH instead of its trace because DoH bears slightly better scale selection property under non-Euclidean affine transformation than LoG. Bay et al. used a Fast-Hessian detector with box filters to approximate DoH. The SURF descriptor is a 64D vector composed by the Harr wavelet responses of the square area around keypoint. SURF has been demonstrated to outperform SIFT on speed, repeatability, distinctiveness, and robustness [8]. It has been used for multispectral satellite image registration [24], seabed recognition based on sonar images [25], and SAR image registration [26–29].

The next procedure after feature extraction is to match the features for correspondences. For tie points, this procedure is unnecessary because they have already matched when extracted. For other features, the correspondences are usually constructed by optimizing certain merit function, such as maximizing the similarity

or minimizing the difference. The warp function can then be retrieved by fitting the obtained correspondences. For correspondences without any mismatches, the retrieval can be easily conducted by fitting them with the least squares (LS). However, for the general registration cases, the initial correspondences often contain mismatches. Therefore, the robust retrieval algorithms which are insensitive to outliers are needed. In many existing literatures on feature-based SAR image registration [15, 16, 26, 27], the random sample consensus (RANSAC) [30] has been widely used and recommended for warp function retrieval. RANSAC conducts the estimation by randomly sampling a minimal sampling set (MSS) to achieve an estimation of the warping, and the entire datasets are then checked on the estimation for a consensus set (CS) of correspondences. These two steps are iterated until the largest CS is achieved [31]. Besides this, the least median squares (LMedS) [32] and the fast least trimmed squares (Fast-LTS) [33] have also been used [4, 34, 35]. There are also some other approaches which use different matching and retrieval algorithms with different features, which can be referred to the related reviewing articles [36–38].

Although lots of approaches have been developed for feature-based SAR image registration, there are still some open problems that have not been perfectly solved yet. In this chapter, we concentrate on two problems, i.e., which feature is more appropriate and which retrieval algorithm performs much better? The first problem is related to the feature operator, which is focused in Sections 2 and 3. We give a detailed evaluation to tie points, Harris corner, SIFT, and SURF in terms of the geometrical invariance of feature and descriptor, extraction and matching speed, localization accuracy, robustness to decorrelation, and flexibility to speckle. SURF is identified to outperform others. Particularly, we find that SURF is flexible to speckle for the close relationship between Fast-Hessian detector and refined Lee speckle filter. SURF is thus more competent for SAR image registration. The second problem is posed in Section 4 with the reason that the widely used RANSAC is found instable for parameter estimation in the registration of an interferometric SAR (InSAR) image pair. The uncertainty arises from its inappropriate loss function and estimation strategy. Based on the scheme of Fast-LTS, an extended Fast-LTS (EF-LTS) is presented for 2D robust parameter estimation. Experiment on InSAR image pair demonstrates that EF-LTS is more stable and robust than RANSAC. It is more appropriate and competent for SAR image registration. Based on these, we recommend fitting the SURF features with EF-LTS to conduct the registration. We further evaluate this scheme in Section 5 by processing the MiniSAR image pair, and the result complies with our expectation. Section 6 concludes the chapter finally.

2. Comparative analysis on the commonly used features for SAR image registration

SAR image is acquired with intensity and phase, which should be transformed into the real one before feature detection by taking the intensity or the logarithmic intensity of the image. Instead of proposing a novel feature for SAR image registration, we identify the appropriate feature from the widely used tie points, Harris corner, SIFT, and SURF by evaluating them on several criteria. In this section, the features will be evaluated on the following six factors, i.e., the geometrical invariance of feature, the extraction speed, the localization accuracy, the geometric invariance of descriptor, the matching speed, and the robustness to decorrelation, while the impact of SAR speckles will be particularly focused and analyzed in Section 3.

2.1 Geometrical invariance of feature

The geometrical invariance of feature refers to which degree of warping a same feature can still be extracted from the warped images by a detector. Cross-correlation (CC) is sensitive to image rotation and scaling, hence the CC-based tie points are only invariant to the following translation transformation:

$$
\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
\tag{1}
$$

where $(x, y, 1)^{\mathrm{T}} \leftrightarrow (x', y', 1)^{\mathrm{T}}$ are the inhomogeneous coordinates of a pair of matching points (the superscript T shows the vector transpose), and t_x and t_y denote the translations in x- and y-direction, respectively. The Harris measure is the following Harris matrix $\mathbf{H}$ describing the neighboring gradient distribution of a point [5]:

$$
\mathbf{H} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}
\tag{2}
$$

where $\langle \cdot \rangle$ denotes the ensemble average; I_x and I_y are the first-order partial derivatives in x- and y-direction, respectively. Then, the response function R of Harris is the weighted sum of the determinant and squared trace of $\mathbf{H}$ [5]:

$$
R = \det(\mathbf{H}) - \kappa(\mathrm{trace}(\mathbf{H}))^2
\tag{3}
$$

where the weight κ is a constant within the interval 0.04–0.06. A pixel is selected as a Harris corner if its response R is beyond a given threshold. It can be easily obtained from (2) that $\mathbf{H}$ is semi-definite Hermitian, which indicates the existence of two nonnegative eigenvalues λ_1 and λ_2. Then (3) can be further formulated as:

$$
R = \lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2)^2
\tag{4}
$$

The Harris response R is only decided by the eigenvalues of $\mathbf{H}$. Any unitary transformation of $\mathbf{H}$ will not influence the extraction of corner. Therefore, Harris corner is invariant to the following Euclidean transformation:

$$
\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
\tag{5}
$$

where θ denotes the rotation. SIFT and SURF were proposed to achieve the scale-invariance further:

$$
\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
\tag{6}
$$

where s is the scale. Theoretically, SIFT and SURF features are not affine-invariant as Harris-Affine and Hessian-Affine features [39]. Nonetheless, the affine frame in Hessian-Affine and Harris-Affine is more sensitive to noise than

scale-invariant detector. For general SAR image application, scale-invariant features such as SIFT and SURF are sufficient.

2.2 Feature extraction speed

The extraction speed is mainly influenced by the computational load of detector. Tie points are identified by traversing all potential offsets to calculate CC. The resulted computational load is heavy. The Harris response R is determined by the determinant and trace of matrix $\mathbf{H}$. The calculation of $\mathbf{H}$ only relates to the first-order derivatives which can be fast achieved. The scale-invariant SIFT and SURF keypoints are extracted by constructing SSP first. SSP is comprised of several octaves and each octave consists of several scale levels further. A scale level is a Gaussian-smoothed image. The nearby two layers are subtracted to calculate DoG, an approximation to LoG. The keypoint is finally identified as the point with extreme value of DoG in a $3 \times 3 \times 3$ neighborhood in the scale space. SIFT detector performs slower than Harris because it extracts the feature in 3D space not in 2D space. Nonetheless, to extract the same number of subpixel features, SIFT detector is faster than CC-based tie points for the latter conducts exhaustive searching. SURF extracts feature based on DoH. Given a point $\mathbf{x} = (x, y)$ in image $\mathbf{I}$ at scale σ, the scale function DoH is obtained by:

$$DoH_{SIFT} = L_{xx}(\mathbf{x}, \sigma)L_{yy}(\mathbf{x}, \sigma) - \left(L_{xy}(\mathbf{x}, \sigma)\right)^2 \tag{7}$$

where $L_{xx}(\mathbf{x}, \sigma)$, $L_{yy}(\mathbf{x}, \sigma)$, and $L_{xy}(\mathbf{x}, \sigma)$ denote the convolution of the Gaussian second-order derivative in x-, y-, and xy-directions with $\mathbf{I}$, respectively.

When applied in 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 displayed in the first row of **Figure 1**. Encouraged by the successful simplification of LoG with DoG in SIFT, Bay et al. devised a Fast-Hessian detector to approximate L_{xx}, L_{xy}, and L_{yy} with box filters D_{xx}, D_{xy}, and D_{yy}, respectively, shown in the second row of **Figure 1**. In [8], Bay et al. indicated that the performance of this approximation is comparable or even better than the original Gaussians. The
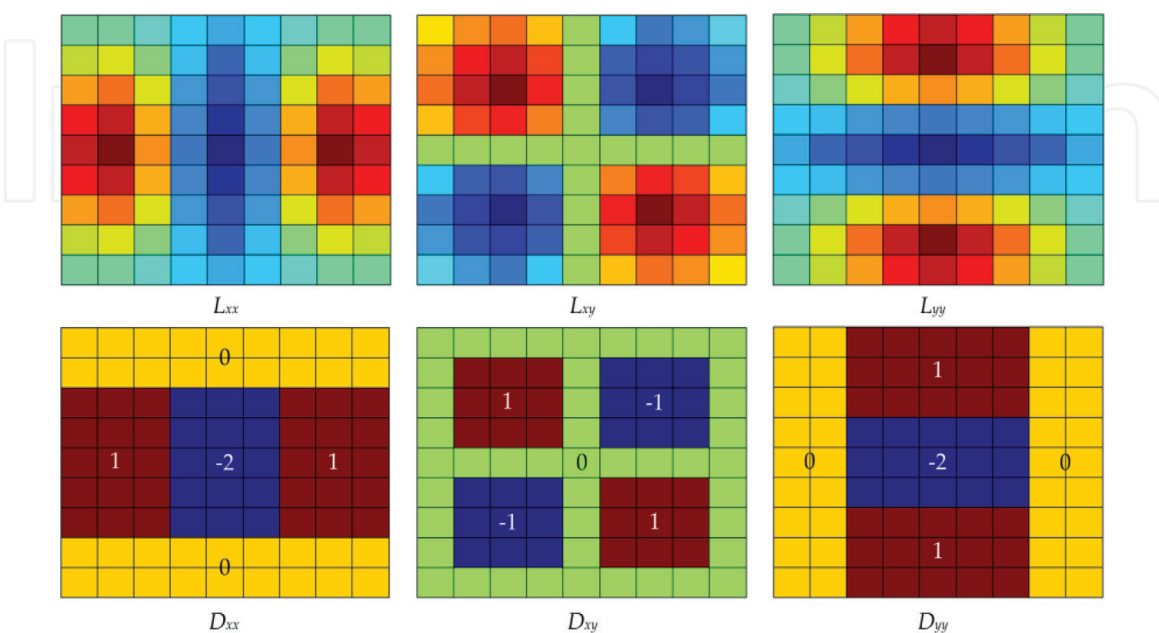

Figure 1.
SIFT discretized and cropped Gaussian second-order partial derivatives in x- (L$_{xx}$), xy- (L$_{xy}$), and y-direction (L$_{yy}$), as well as their corresponding SURF box filter approximations D$_{xx}$, D$_{xy}$, and D$_{yy}$, respectively.

approximation makes pixels in certain window have the same weight. The convolutions can be then calculated at very low computational cost by using the integral image. Therefore, instead of iteratively reducing the image size and using the cascade filtering, SSP in SURF is built by simply up-scaling the box filters without changing the size of the image. The use of integral image enables the convolutions independent of the filter size and scale.

2.3 Localization accuracy of feature

Image registration accuracy is closely determined by the localization accuracy of feature. Tie points achieve subpixel accuracy by oversampling the image patches [40] or CC obtained in coarse registration [41]. Higher sampling rate indicates higher accuracy, but it also signifies larger data sets, heavier computational load, and more severe aliasing. Keypoint in SIFT and SURF is first located as the extrema using the non-maximum suppression technique, and is then refined to subpixel and sub-scale accuracy by Taylor fitting a 3D quadratic to the scale function DoG (for SIFT) or the approximated DoH (for SURF) in the scale space [42]:

$$f(\mathbf{X}) = f(\mathbf{X}_0) + \left(\frac{\partial f}{\partial \mathbf{X}}(\mathbf{X}_0)\right)^{\mathrm{T}} \Delta\mathbf{X} + \frac{1}{2}\Delta\mathbf{X}^{\mathrm{T}}\left(\frac{\partial^2 f}{\partial \mathbf{x}^2}(\mathbf{X}_0)\right)\Delta\mathbf{X}. \tag{8}$$

Therefore, SIFT and SURF can obtain 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. For high accurate SAR image registration, we should further evaluate the features carefully, and this will be detailed in Section 3.4.

2.4 Geometrical invariance of descriptor

Feature descriptor is usually a vector depicting the neighboring information of a feature. It plays a key role in feature matching. The descriptor's geometrical invariance determines the degree of warping to which features can still be successfully matched. Harris corner and tie points have no descriptor. From feature matching point of view, however, they both adopt template matching by selecting the image square centered around the feature as descriptor, which is only invariant to translation. Thus, tie points and Harris corner can be successfully matched only under weak warping. SIFT and SURF descriptors enable a good compromise between feature complexity and the robustness to commonly occurring deformation such as weak affine transformation [7, 8, 43]:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x \cos\theta & -s_y \sin\theta & t_x \\ s_x \sin\theta & s_y \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{9}$$

where s_x and s_y denote the scales in directions x and y, respectively. Robust matching across a substantial range of affine distortion and change in 3D viewpoint can hence be achieved.

2.5 Matching speed of feature

Feature matching is usually conducted based on certain merit function of the descriptors. In feature-based SAR image registration, the merit function is to

maximize the similarity (such as CC [4]) or minimize the differences (such as Euclidean distance [7, 8]). A correspondence is detected if it can optimize the merit function. For SIFT and SURF, the merit of an optimal correspondence has also to be certain times larger than the second optimal merit. Matching speed is mainly determined by the calculation of merit. For tie points and Harris corner, the merit function is the maximum of CC, which can be obtained on complex data or magnitude data [44], referring to coherent CC or incoherent CC, respectively. The registration accuracy attained by coherent CC is much higher than that by incoherent CC [45]. If D_1 and D_2 are the image patches, respectively, centered at an initial match, the coherent CC is calculated as

$$CC(D_1, D_2) = \frac{\left| \sum_{i=1}^{N} \sum_{j=1}^{N} (D_1(i,j) - \mu_1)(D_2(i,j) - \mu_2)^* \right|}{\sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} |(D_1(i,j) - \mu_1)|^2 \sum_{i=1}^{N} \sum_{j=1}^{N} |(D_2(i,j) - \mu_2)|^2}} \quad (10)$$

where N is the size of the image patch, μ_1 and μ_1 denote the means of D_1 and D_2, respectively. Equation (10) requires about $10N^2$ operations including $7N^2$ additions and $3N^2$ multiplications.

The merit function in SIFT and SURF is the minimum of the Euclidean distance. If D_3 and D_4 are the descriptors of an initial match, respectively, the distance can be calculated by

$$Dist(D_3, D_4) = \sum_{i=1}^{L} |D_3(i) - D_4(i)|^2 \quad (11)$$

where L is the length of descriptor. Equation (11) requires $3L$ operations including $2L$ additions and L multiplications. For SURF, Bay et al. [8] found that the sign of Laplacian can be further used to distinguish the feature from its background for fast indexing during matching stage. The merit will not be computed unless the initial match has the same sign. Hence, under the assumption of equal probability distribution for sign of Laplacian, the merit computation in SURF requires $1.5L$ operations. Taking the descriptor lengths L for SIFT and SURF being 128 and 64 into consideration, then (11) involves in 384 and 96 operations for SIFT and SURF, respectively. Hence, SURF is four times faster than SIFT on feature matching. To achieve the same efficiency as SIFT or SURF, the equivalent patch size N for tie points and Harris corner should be about 6 or 3, respectively. This may lead to biased CC estimation thus bad feature localization and matching due to the insufficient sampling.

2.6 Robustness to decorrelation

SAR decorrelation sources can be classified into two categories, i.e., the geometrical warping and radiometric warping. Geometrical warping will lead to decorrelation and influence the CC-based feature matching, which relates to the geometrical invariance of feature discussed above. Here, we focus on the radiometric warping-induced decorrelation. Such decorrelation is resulted because CC is only invariant to affine changes in scattering. Target scattering in microwave band is sensitive to frequency, bandwidth, and polarization. All these introduce a complex nonlinear radiometric warping, which degrades SAR information and aggravates image registration by impacting the localization of tie points. The localization accuracy of tie points is measured by the error standard deviation σ_L [45]:

$$\sigma_L = \sqrt{\frac{3}{2N^2}} \frac{\sqrt{1-\gamma^2}}{\pi\gamma} osr^{3/2} \tag{12}$$

where γ is CC, N is the size of the tie patches, and *osr* is the oversampling rate. Localization accuracy directly relates to CC: higher coherence means higher localization accuracy, while higher decorrelation indicates worse localization accuracy and worse registration accuracy. It is known that one can approximate a nonlinear function with a series of linear functions, so a nice method to improve the robustness to decorrelation is to use smaller image patches, but this will also result in worse localization accuracy through N in (12). Thus, tie points are not robust to decorrelation. Similarly, the influence of decorrelation on CC-based matching of Harris corners is also unavoidable. However, Harris, SIFT, and SURF locate feature based on geometrical texture instead of correlation. This will reduce the influence of decorrelation. The matching of SIFT and SURF features is based on local descriptors which are invariant to affine changes in scattering. SIFT and SURF features are thus more robust to decorrelation.

3. Impact of SAR speckles on accurate feature extraction

SAR image is acquired by actively measuring and coherently processing the electromagnetic scattering of target. The interference of scatterings from scatterers within each resolution cell produces a pixel-to-pixel variation in image intensity and results in the so-called speckle. In this section, we first conduct a qualitative evaluation on the flexibility of existing features to speckles. An experimental evaluation of the identified feature is then conducted and some necessary improvements are developed for high accurate SAR image registration.

3.1 Flexibility to image speckling

For CC-based tie points, the assumption that the scattering is locally stationary and ergodic may not be tenable in the existence of speckles. As a result, the correlation estimation as well as the localization and matching of the feature will be biased. For the geometrical texture-based detectors such as Harris, SIFT, and SURF, speckles may lead to false texture and high MFAR. To achieve stable features from the speckle-contaminated SAR image, a conceivable method is to suppress speckle beforehand. Schwind et al. [15] suggested adopting the ISEF filter, but they indicated that ISEF filter and any other filter may slightly affect feature localization and registration quality. Hence, a better strategy is to conduct speckle suppression while feature extraction, i.e., the detector should be flexible to speckling.

Harris detector obtains features using the first-order image derivatives which are not robust to speckles. As a result, Harris detector may extract many features, but most of the extracted features are speckles with only a few correct matches. This influence has been also observed by Schwind et al. [15] when using SIFT to SAR: only very few matches are constructed at the first octave of SSP although with extensive number of extractable features, and the matches from this octave have the highest MFAR of all the octaves. The first scale octave refers to the original or double-sized images which are of the highest resolution and the largest number of extractable keypoints. The highest MFAR at this octave clearly indicates the bad flexibility of SIFT to speckles, while the lower MFAR at higher octaves is just due to the fact that larger image smoothing reduces the speckle. Different from SIFT, SURF can deal with speckle very well because of the relationship between Fast-Hessian detector and refined Lee speckle filter.

3.2 Refined Lee speckle filter

An ideal speckle filter should adaptively smooth speckle, retain the sharpness of boundaries and edges, and preserve the subtle but distinguishable details. The most widely used boxcar filter replaces a pixel with the mean of its windowed neighborhood. This filter can be easily implemented and works very well in homogeneous area, but will degrade spatial resolution in inhomogeneous area due to the indiscriminate averaging [46]. To solve this, many filtering techniques have been proposed. The refined Lee speckle filter is just such a filter which uses the local statistics to suppress speckles without degrading image. To identify pixels with the similar texture, Lee devised the eight non-square edge-aligned windows, as shown in **Figure 2**. In the course of filtering, one of the windows is matched to calculate local statistics based on edge direction, and the minimum mean square algorithm is then adopted for filtering. As a result, this filter can effectively reduce the speckle without degrading the edge [46].

3.3 Relationship between Fast-Hessian detector and refined Lee filter

As mentioned previously, SURF extracts features based on the box filter displayed in **Figure 1**. Box filter not only speeds up feature extraction, but also enables SURF to extract features while reducing speckles. In D_{xx} of **Figure 1**, we average the pixels using a 5×3 window first, and then extract the vertical edge by the second-order image partial derivative in x-direction with convolution template $[1 -2\ 1]$. This is equivalent to filter speckles with Lee's windows **Figure 2(a)** and **(e)**. Similarly, D_{yy} denotes that we also filter the pixels using a 5×3 window first, but then extract the horizontal edge using the second-order image partial derivative in y-direction with convolution template $[1 -2\ 1]^{\mathrm{T}}$. This is equivalent to filter speckle with Lee's non-square windows **Figure 2(c)** and **(g)**. D_{xy} shows that we use a 3×3 window and extract the 135° edge feature by the second-order image partial derivative in negative xy-direction with the convolution template $[1 -1; -1, 1]$. This is equivalent to filter speckle with windows **Figure 2(d)** and **(h)**. Likewise, $-D_{xy}$ gives that we also use a 3×3 window but extract the 45° edge by the second order image partial derivative in positive xy-direction with convolution template $[-1\ 1; 1, -1]$. This is equivalent to filter with windows **Figure 2(b)** and **(f)**. Instead of
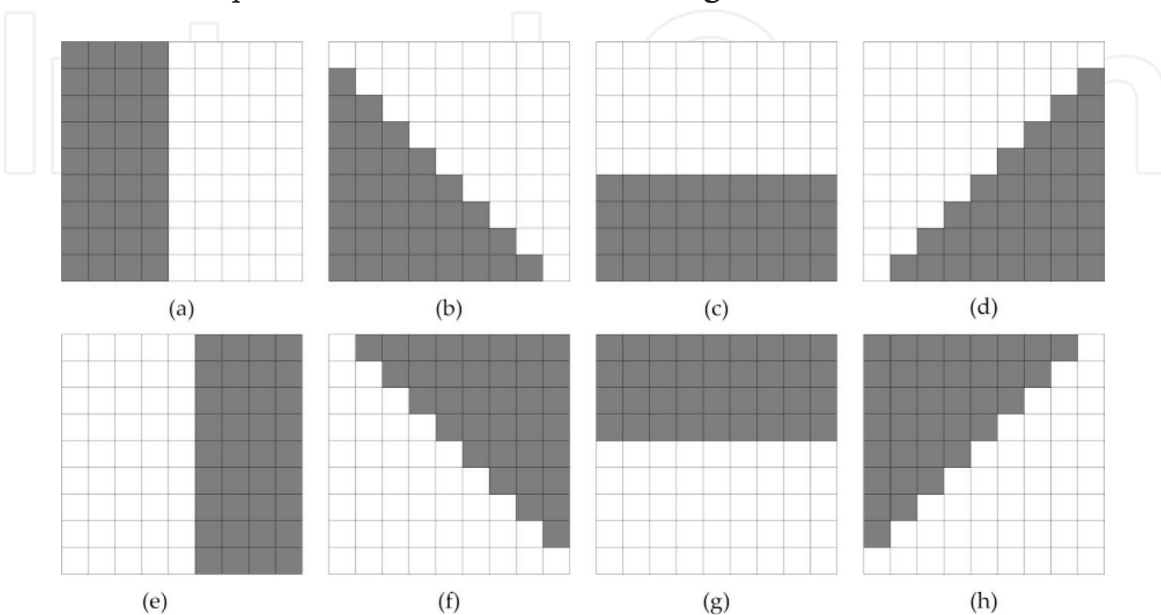

Figure 2.
Edge-aligned windows used in refined Lee filter to decide the local texture, where windows (a) and (e) are used for vertical edge, (c) and (g) for horizontal edge, (b) and (f) for 135° edge, and (d) and (h) for 45° edge. The pixels in white are used for filtering computation.

selecting the optimal edge to calculate local statistics, the four edge features are combined to a new feature in SURF by:

$$DoH_{SURF} = D_{xx}D_{yy} + (0.9D_{xy})(-0.9D_{xy}) \qquad (13)$$

which corresponds to DoH in (7), where the constant 0.9 is used to balance the expression for the Hessian's determinant. Then, SSP in SURF just indicates that we adopt a series of box filters of different size to filter speckles and extract features of different scales. Hence, SURF is very flexible to deal with speckle.

3.4 Evaluation of SURF for SAR image subpixel registration

As listed in **Table 1**, according to the comparative analysis in Sections 2 and 3.1 on several criteria, we can obtain that for the general registration of SAR images

- SURF outperforms others in terms of the considered criteria.

- SIFT is applicable when no strict requirement for speed.

- Harris may be appropriate for coarse registration.

- Tie points are fit for images with slight distortion and weak decorrelation and require heavy computation load.

From these, we can see that SURF is more appropriate and competent for general SAR image registration. Nevertheless, SAR applications, like DEM retrieval and deformation estimation usually impose a strict requirement for registration accuracy. To ensure an acceptable result, the registration accuracy should be subpixel. To evaluate the capability of SURF for subpixel image registration, we devise a comparative experiment on some contrived SAR image pairs. **Figure 3** shows a SAR image of Enta Volcano acquired by SIR-C/X-SAR. We treat this image as the master and transform it to model an affine geometrical warp for the slave image:

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*	Pixel	Subpixel	Subpixel
Geometrical invariance of feature descriptor	Translation	Translation	Affine transform	Affine transform
Feature matching speed	Slow	Slow	Fast	Faster
Robustness to decorrelation	Worse	Bad	Good	Good
Flexibility to image speckle	Good	Bad	Bad	Better

Determined by the sampling rate.

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

$$\begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x_m \\ y_m \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ 1 \end{bmatrix} \quad (14)$$

where $(x, y, 1)^{\mathrm{T}}$ are the homogenous image coordinates, subscripts s and m denote the slave and master images, respectively. $\mathbf{A}$ is an affine matrix composed by parameters a, b, c, and d, as well as two translations t_x and t_y. Bay et al. devised two versions of Fast-Hessian detectors for SURF. The one initializes SSP by using 9×9 box filter to the original image is denoted as *FH-9(-1)*, while the one initializes SSP by using 15×15 box filter to double-sized image (also with doubled sampling step) is denoted as *FH-15(-2)*. *FH-15(-2)* has been shown to be better than *FH-9(-1)* on repeatability [8]. We use the two detectors to extract point correspondences, respectively, based on which the robust EF-LTS (will be presented in Section 4) is then used to retrieve the warp matrix. To compare the two SURF detectors for SAR image registration, we consider four criteria, i.e., the average transfer error (*ATE*), *MFAR*, the number of correct matches, and the warp matrix estimation error (*WMEE*). *ATE* measures the appropriateness of the extracted features to the achieved warp parameters:

$$ATE = \frac{1}{N} \sum_{i=1}^{N} \left\| \begin{bmatrix} x_{si} \\ y_{si} \\ 1 \end{bmatrix} - \mathbf{A_r} \begin{bmatrix} x_{mi} \\ y_{mi} \\ 1 \end{bmatrix} \right\| \quad (15)$$

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

$$\left\| \begin{bmatrix} x_{si} \\ y_{si} \\ 1 \end{bmatrix} - \mathbf{A} \begin{bmatrix} x_{mi} \\ y_{mi} \\ 1 \end{bmatrix} \right\| < threshold \begin{cases} \text{True} & (x_{si}, y_{si}) \leftrightarrow (x_{mi}, y_{mi}) \text{ is a correct match} \\ \text{False} & (x_{si}, y_{si}) \leftrightarrow (x_{mi}, y_{mi}) \text{ is a mismatch} \end{cases}$$
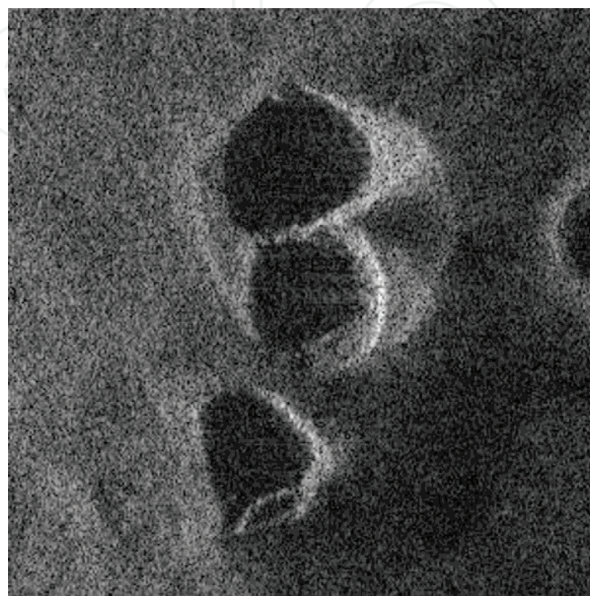
$$(16)$$

Figure 3.
SAR image of Enta Volcano taken by SIR-C/X-SAR (300 × 300).

where **A** is the true warp matrix. The *threshold* is chosen as 5 pixels, i.e., a correspondence is identified as a mismatch if the transfer error is larger than 5 pixels in any image direction.

MFAR, also called 1-precision [10], is defined as:

$$MFAR = \frac{\#matches - \#correct\ matches}{\#matches} \qquad (17)$$

where "#" denotes "the number of." *MFAR* is just the rate of mismatches, which is related to image speckling as well as the radiometric and geometrical warping. It can be used together with *#correct matches* to evaluate the robustness of a detector to speckles on SAR image pair with controlled radiometric and geometrical warping.

WMEE is used to evaluate the consistency of the retrieved warp matrix and its true value:

$$WMEE = \|\mathbf{A} - \mathbf{A_r}\|_F \qquad (18)$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

We evaluate the two SURF detectors on four image pairs with different transformations, the retrieved warp matrix parameters, *ATE*, correct match number, *MFAR*, and *WMEE* are listed in **Table 2**. It shows that *FH*-15(-2) can extract more correct matches with lower *MFAR* than *FH*-9(-1). This validates the robustness of SURF to speckling because *FH*-15(-2) performs the feature extraction on the double-sized image with much serious speckle. *ATE* of *FH*-15(-2) is smaller than that of *FH*-9(-1) except on the first image pair. On all the four pairs, the features extracted by *FH*-15(-2) can obtain subpixel estimation in both image directions, but *FH*-9(-1) obtains this only on the first pair. Therefore, *FH*-15(-2) features are more consistent with the retrieval parameters. This also signifies that *FH*-15(-2) can attains lower *MFAR* than *FH*-9(-1) because parameter estimation in EF-LTS is related to the outlier percentage in data. This will be detailed in Section 4. As on *WMEE*, the two detectors perform equally, *FH*-15(-2) does not improve the registration accuracy on all the four pairs as we expected, and there is still clear inconsistency between the retrieved warp matrix and the true value. The reason lies in that the sampling step is also doubled when *FH*-15(-2) doubles the image. This makes sampling being still conducted on the equivalently same pixel position rather than the subpixel image position. For instance, let (x_0, y_0) be a sampled pixel in the original image, the corresponding position in doubled image is $(2x_0, 2y_0)$. The doubled step then makes this pixel position be still sampled instead of $(2x_0 \pm 1, 2y_0 \pm 1)$, while the latter corresponds to the subpixel position $(x_0 \pm 0.5, y_0 \pm 0.5)$ in the original image and positively contributes to the subpixel registration. Based on this, we suggest initializing SSP by using 9×9 box filter to the oversampled image but with unchanged sampling, we denote this detector as *FH*-9(-*Fs*), *Fs* denotes the sampling rate. To avoid nonlinear aliasing, the linear interpolator such as bilinear interpolator is used to conduct the sampling. **Table 2** further summarizes the registration results based on *FH*-9(-2) to *FH*-9(-5) detector. Comparing with *FH*-9(-1) and *FH*-15(-2), the correct match number, *ATE*, *MFAR*, and *WMEE* of *FH*-9(-2) are all clearly improved. As oversampling rate increases from 2 to 5, the registration accuracy is also improved for more correspondences of higher localization accuracy are identified. All these make the high accurate SAR image registration possible. In view of the fact that oversampling will increase dataset and computational load, for high accuracy registration we recommend oversampling the image three or four times so as to achieve the compromise among accuracy, robustness, and computational complexity.

Detectors	Estimated affine warping parameters						Correct match number 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)	0.7164	0.0415	−0.0481	0.8059	2.3151	3.6269	42 (0.1923)	(0.7109, 0.8770)	1.3725
FH-9(-1)	0.7195	0.0425	−0.0347	0.8067	2.0444	1.3480	22 (0.2414)	(0.5887, 0.7854)	1.1070
FH-9(-2)	0.7186	0.0458	−0.0395	0.8085	1.3565	2.2052	73 (0.1300)	(0.6219, 0.6602)	0.3949
FH-9(-3)	0.7192	0.0450	−0.0403	0.8088	1.4752	2.3425	129 (0.1164)	(0.3001, 0.4602)	0.2321
FH-9(-4)	0.7181	0.0453	−0.0402	0.8093	1.6070	2.1746	188 (0.1754)	(0.2580, 0.3790)	0.2439
FH-9(-5)	0.7186	0.0457	−0.0398	0.8094	1.4895	2.1085	176 (0.1619)	(0.2819, 0.3874)	0.3596
True value	0.9361	0.1889	−0.1617	1.0938	−10.5000	−3.4000	—	—	—
FH-15(-2)	0.9370	0.1887	−0.1576	1.0908	−10.5603	−3.7546	55 (0.0678)	(0.6040, 0.7075)	0.3598
FH-9(-1)	0.9298	0.1909	−0.1603	1.0868	−9.9304	−2.2059	25 (0.2188)	(0.7949, 1.2405)	1.3231
FH-9(-2)	0.9352	0.1898	−0.1618	1.0940	−10.4452	−3.4432	170 (0.0449)	(0.3821, 0.5200)	0.0698
FH-9(-3)	0.9361	0.1890	−0.1613	1.0937	−10.4329	−3.4817	419 (0.0141)	(0.2267, 0.3080)	0.1058
FH-9(-4)	0.9361	0.1890	−0.1617	1.0937	−10.4252	−3.4490	735 (0.0252)	(0.1703, 0.2143)	0.0894
FH-9(-5)	0.9360	0.1890	−0.1616	1.0938	−10.4227	−3.4476	893 (0.0262)	(0.1601, 0.2273)	0.0908
True value	1.1365	0.1036	−0.0894	1.3159	−2.6000	5.4000	—	—	—
FH-15(-2)	1.1387	0.0984	−0.0736	1.3238	−2.2175	1.7098	47 (0.0408)	(0.7131, 0.9156)	3.7101
FH-9(-1)	1.1402	0.1055	−0.0805	1.3160	−3.6578	3.5156	29 (0.1212)	(1.0153, 0.9129)	2.1610
FH-9(-2)	1.1361	0.1038	−0.0897	1.3160	−2.3833	5.6308	157 (0.0427)	(0.3856, 0.4829)	0.3166
FH-9(-3)	1.1363	0.1038	−0.0895	1.3165	−2.4408	5.4805	476 (0.0206)	(0.1902, 0.3197)	0.1784
FH-9(-4)	1.1365	0.1037	−0.0894	1.3159	−2.4575	5.5336	983 (0.0180)	(0.1616, 0.2378)	0.1954
FH-9(-5)	1.1363	0.1037	−0.0894	1.3160	−2.4582	5.5270	1293 (0.0300)	(0.1432, 0.2119)	0.1903

Detectors	Estimated affine warping parameters						Correct match number and *MFAR*	*ATE*	*WMEE*
	a	*b*	*c*	*d*	t_x	t_y			
True value	1.2079	0.0777	−0.0718	1.3077	−5.3000	1.5000	—	—	—
FH-15(-2)	1.2033	0.0744	−0.0753	1.3054	−4.2454	2.4055	52 (0.0545)	(0.7247, 0.7616)	1.3900
FH-9(-1)	1.1959	0.0695	−0.0650	1.3079	−1.8954	0.0939	24 (0.0769)	(0.9570, 1.1486)	3.6836
FH-9(-2)	1.2075	0.0778	−0.0735	1.3066	−5.0414	2.0074	172 (0.0227)	(0.3858, 0.4182)	0.5695
FH-9(-3)	1.2076	0.0766	−0.0719	1.3077	−5.0751	1.6740	514 (0.0172)	(0.2207, 0.3552)	0.2844
FH-9(-4)	1.2078	0.0777	−0.0718	1.3077	−5.1181	1.6590	1052 (0.0177)	(0.1506, 0.2289)	0.2416
FH-9(-5)	1.2077	0.0778	−0.0719	1.3078	−5.1297	1.6397	1451 (0.0176)	(0.1343, 0.1998)	0.2203

Table 2.
Evaluation of SURF Fast-Hessian detectors on four controlled SAR image pairs.

4. Appropriate retrieval algorithm for SAR image registration

The next procedure after feature extraction is to retrieve the warp function from the attained correspondences. Due to the influences of spatial/temporal decorrelation, system noise, and environmental interference, or the non-robustness in the depiction and matching of features, there are always mismatches in the constructed correspondences. It is difficult to get *a priori* information to remove them beforehand. To accurately retrieve parameters from these error-prone correspondences, some robust outlier-insensitive algorithms are necessary.

Furthermore, unlike the pinhole imaging of optical camera, SAR acquires the imagery using a slant-range geometry which cannot be modeled as a central projection [47]. As a result, the warp model between SAR images is dependent on the system parameter, imaging geometry, and target relief, and we cannot adopt a global homography or essential matrix to model the geometrical warping then. Nevertheless, when the system parameter and imaging geometry are fixed and the area-of-interest has gentle topography, we can conventionally approximate the warp function as a low-order polynomial [48]. This indicates our strategy in the retrieval of registration parameters, to focus on the global registration instead of local discontentment.

4.1 Evaluation of RANSAC for SAR image registration

RANSAC [30] has been widely used in feature-based SAR image registrations for parameter retrieval [15, 16, 26, 27]. Unlike LS which uses all the available data to estimate parameters, RANSAC conducts the estimation using a few-to-many strategy or a local-to-global strategy. A MSS is randomly sampled from the constructed correspondences to achieve an estimation of the warp function firstly. The cardinality of MSS, i.e., the smallest sufficiency to determine the warp parameters, is just related to the degree of freedom (DoF) of the warp function. For example, the cardinality will be 3 for affine transformation of 6 DOFs. The entire dataset are then checked for those correspondences consistent with the retrieved warping to construct a larger CS. These two steps are repeated until the largest CS is finally achieved for parameter estimation. This local-to-global strategy is tenable only if any MSS of inliers can generate the "true value" of warp parameters [31]. But it is often hard to keep this in real registration due to the unavoidable noise and local distortion, i.e., a different estimation of parameters will be achieved from a different MSS configuration of inliers. This uncertainty is even more severe in SAR image registration because SAR warping varies from pixel to pixel and the low-order polynomial approximation only accounts for global registration instead of local contentment. The local-to-global strategy may then magnify the local distortion, aggravate the estimation uncertainty, and damnify the global registration accuracy although a largest CS is identified. To demonstrate this, we devise an experiment to coregister a spaceborne InSAR image pair as shown in **Figure 4(a)** and **(b)**. The two images are acquired by RadarSat-2 on May 4 and 28, 2008, respectively. The scene is within South Phoenix, AZ, USA with some buildings and vegetable lands. We first use *FH*-9(-1) to construct SURF feature correspondences, and then adopt RANSAC to retrieve the affine warp parameters. To evaluate the estimation certainty, we execute RANSAC 100 times and based on the obtained parameters of each execution, we coregister the complex image pair to calculate the three-look coherent CC and spectral SNR. CC measures the consistency, while spectral SNR, the ratio between the maximum entry and the sum of other entries in the spectrum, reflects the clarity of the interferogram fringe [49]. **Figure 5** displays the affine

parameters a, b, c, d, t_x, and t_y as well as CC and SNR obtained in each execution. **Table 3** further displays the mean and standard deviation of the parameters, CC, and SNR. RANSAC cannot obtain a stable registration because the retrieval parameters vary with executions, even for executions with the same cardinality of CS achieved. **Figure 6** shows the retrieval parameters, CC and SNR for 48 executions with the same cardinality. We can still find the estimation uncertainty. This reveals that the attained inliers which compose the final CS are actually different although the same cardinality. Otherwise, the parameters would be the same for each execution because they are retrieved by just LS fitting the inliers.

The uncertainty of RANSAC in SAR image registration just comes from its retrieval strategy and loss function. To achieve a stable registration for SAR images, a feasible improvement is to estimate the parameters with more correspondences to reflect the true support than just a MSS, and to apply an appropriate loss function. This leads us another direction to the robust parameter regression.
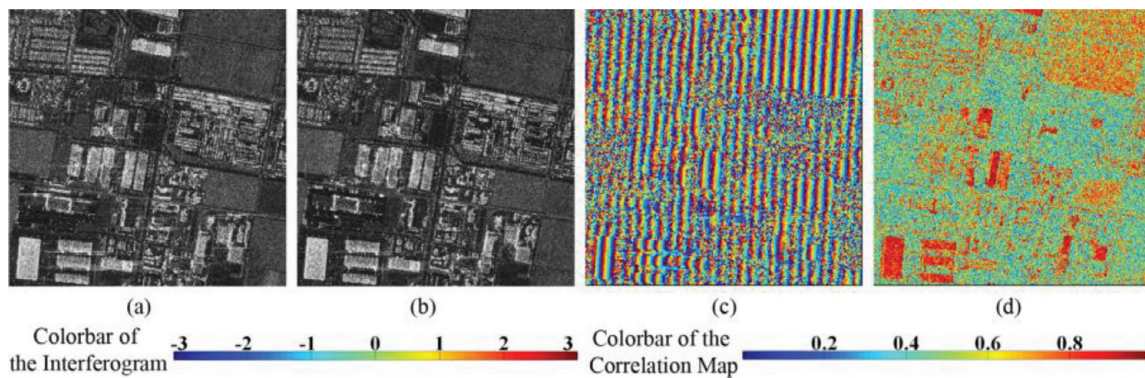

Figure 4.
Registration of InSAR image pair from RadarSat-2. (a) Master image, (b) slave image, and the final (c) interferogram and (d) correlation map based on the registration parameters estimated by EF-LTS.
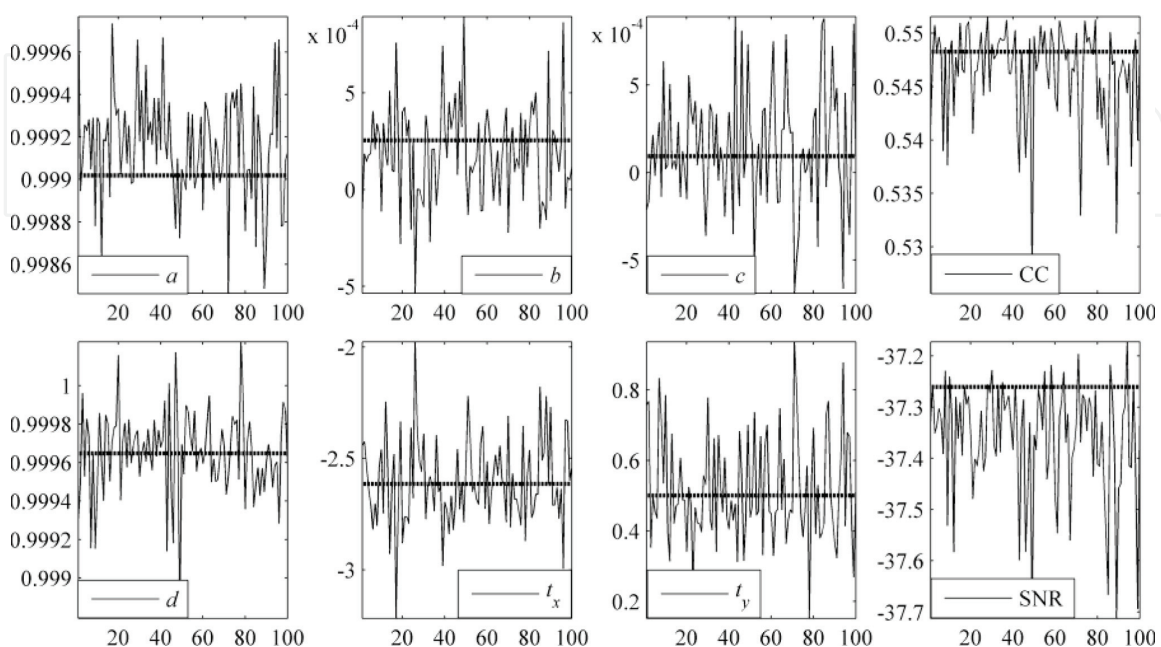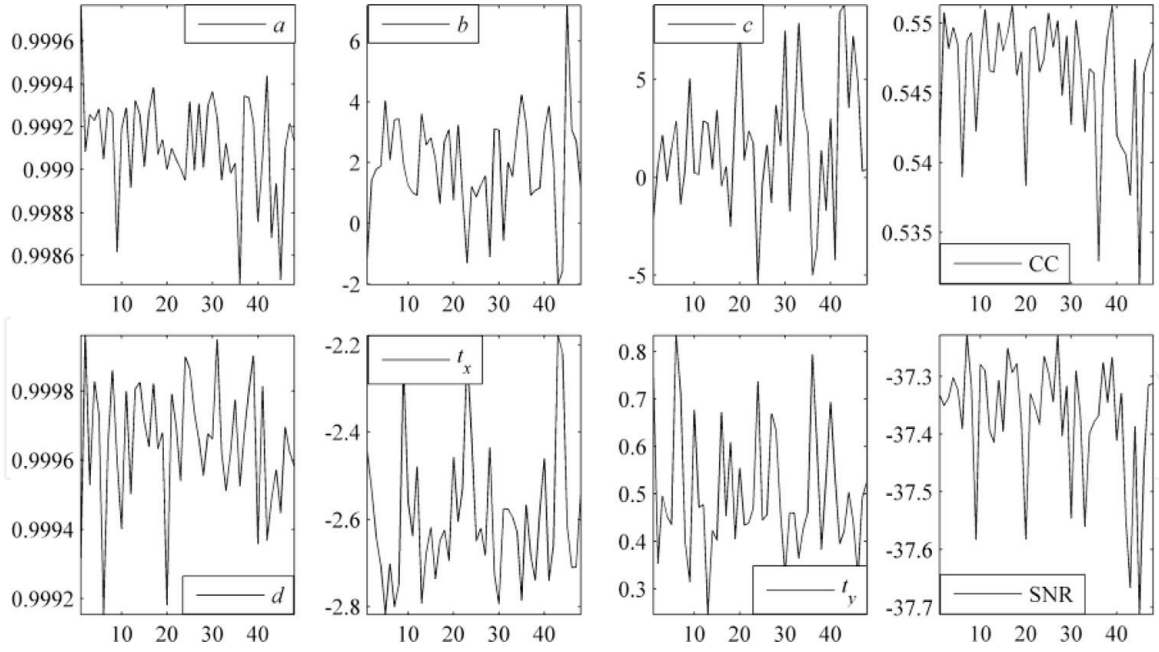

Figure 5.
Registration parameters, cross-correlation (CC), and spectral SNR obtained by 100 executions of (thin line) RANSAC and (thick line) EF-LTS on the image pair of RadarSat-2.

Algorithm	a		b		c		CC	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
RANSAC	0.9992	2.5228×10^{-4}	1.8796×10^{-4}	2.4879×10^{-4}	1.3770×10^{-4}	3.3408×10^{-4}	0.5462	0.0046
EF-LTS	0.9990	0.0000	2.5485×10^{-4}	0.0000	9.2440×10^{-5}	0.0000	0.5483	0.0000
Algorithm	d		t_x		t_y		SNR (dB)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
RANSAC	0.9996	2.2264×10^{-4}	-2.6068	0.1955	0.5133	0.1509	-37.36	0.1126
EF-LTS	0.9996	0.0000	-2.6151	0.0000	0.5008	0.0000	-37.26	0.0000

Table 3.
Mean and standard deviation of the registration parameters, cross-correlation (CC), and spectral SNR obtained by RANSAC and EF-LTS on RadarSat-2 InSAR images.

Figure 6.
Registration parameters, cross-correlation (CC), and spectral SNR obtained by 48 executions of RANSAC on RadarSat-2 InSAR images with the same CS cardinality.

4.2 Fast-LTS

The widely used LS is now being criticized more and more for lack of robustness. To tackle with this, some robust regression approaches were developed, like LMedS [32] and the least trimmed squares (LTS) [50]. LMedS implements the regression by minimizing the median of residual squares. This makes LMedS so robust that it can still obtain a reasonable estimation even if 50% of the dataset are outliers. So the breakdown point of LMedS is as high as 50%. LTS is a modification of LS with the same breakpoint as LMedS. It also fits the linear model:

$$y_i = \mathbf{X}_i^{\mathrm{T}} \boldsymbol{\theta} + e_i, \quad i = 1, ..., n \tag{19}$$

where $\mathbf{X}_i = [x_{i1}, x_{i2}, ..., x_{ip}]^{\mathrm{T}}$ denotes the explanatory variable, y_i denotes the response variable, $\boldsymbol{\theta} = [\theta_1, \theta_2, ..., \theta_p]^{\mathrm{T}}$ indicates the unknown parameter to be retrieved, e_i is the error term, n is the sample size, and p is the dimension of $\mathbf{X}_i$. The loss function of LTS is:

$$Q := \text{Minimize} \sum_{i=1}^{h} (\mathbf{r}^2)_i \text{ with } \mathbf{r} = [r_1, r_2, \cdots, r_n]^{\mathrm{T}} \text{ and } r_i = y_i - \mathbf{X}_i^{\mathrm{T}} \boldsymbol{\theta} \tag{20}$$

where $(\mathbf{r}^2)_i$ denotes the ith element of the ordered squared residuals $(\mathbf{r}^2)_1 \leq \cdots \leq (\mathbf{r}^2)_i \leq \cdots \leq (\mathbf{r}^2)_n$, and h is termed as the trimming constant. LTS conducts regression by LS fitting the h-subset to minimize the squared residuals. Compared with LMedS, the statistical efficiency of LTS is much better and the loss function is much smoother [33]. Nevertheless, the deficiency of LTS is the large computation when processing the big data. To accelerate it, Rousseeuw and Van Driessen [33] developed a Fast-LTS, which can efficiently deal with a sample size as large as tens of thousands or even larger. The core of Fast-LTS is a concentration step (C-step), which is designed to achieve a better estimation from an old h-subset $\mathbf{H_{old}}$ [33]:

Algorithm 1: C-step

Step 11. Compute regression parameters $\boldsymbol{\theta}_{\mathbf{old}}$ by LS fitting $\mathbf{H}_{\mathbf{old}}$.

Step 12. Calculate residuals $\mathbf{r}_{\mathbf{old}}$ based on $\boldsymbol{\theta}_{\mathbf{old}}$. Ascendingly sort squared residuals $\mathbf{r}_{\mathbf{old}}{}^2$ for a permutation $\boldsymbol{\pi}$ of the set such that
$(\mathbf{r}_{\mathbf{old}}{}^2)_{\pi(1)} \leq \cdots \leq (\mathbf{r}_{\mathbf{old}}{}^2)_{\pi(h)} \leq \cdots \leq (\mathbf{r}_{\mathbf{old}}{}^2)_{\pi(n)}$.

Step 13. Construct a new h-subset $\mathbf{H}_{\mathbf{new}} = \{\boldsymbol{\pi}(1), \boldsymbol{\pi}(2), ..., \boldsymbol{\pi}(h)\}$ and obtain the new parameters $\boldsymbol{\theta}_{\mathbf{new}}$ by LS fitting $\mathbf{H}_{\mathbf{new}}$.

It has been proved that Q of parameters $\boldsymbol{\theta}_{\mathbf{new}}$ is always no larger than that of parameters $\boldsymbol{\theta}_{\mathbf{old}}$ [33]. Therefore, an improved estimation of parameters can be achieved after an execution of C-step, and a converged Q will be obtained after only a few C-steps. Thus Fast-LTS conducts estimation as follows [33]:

Algorithm 2: Fast-LTS

Step 21. Randomly generate a p-subset as parameter set $\boldsymbol{\theta}_0$. Calculate n residuals $\mathbf{r}_0$ based on $\boldsymbol{\theta}_0$ to achieve an initial h-subset $\mathbf{H}_0 = \{\boldsymbol{\pi}(1), \boldsymbol{\pi}(2), ..., \boldsymbol{\pi}(h)\}$ such that $(\mathbf{r}_0{}^2)_{\pi(1)} \leq \cdots \leq (\mathbf{r}_0{}^2)_{\pi(h)} \leq \cdots \leq (\mathbf{r}_0{}^2)_{\pi(n)}$. Update $\mathbf{H}_0$ by carrying out two C-steps on $\mathbf{H}_0$. Repeat above procedures 500 times.

Step 22. Implement C-steps on the 10 $\mathbf{H}_0$ with the lowest 10 Q until convergence. Then the solution that creates the lowest Q is identified as the final estimation $\boldsymbol{\theta}$.

The trimming constant h is set between $[(n + p + 1)/2)$ ($[x]$ denotes the smallest integer larger than x) and n. The breakdown value of Fast-LTS is $(n - h + 1)/n$. A nested extension approach should be adopted to enable an efficient estimation when n is larger [33].

4.3 EF-LTS for SAR image registration

Fast-LTS is appropriate for 1D linear regression formulated in (19). However, for SAR image registration, what we need to do is to fit a 2D polynomial regression

$$
\begin{cases}
x_{si} = \sum\limits_{j=0}^{N} \sum\limits_{k=0}^{N-j} a_{jk} x_{mi}^{j} y_{mi}^{k} + \varsigma_i \\
y_{si} = \sum\limits_{j=0}^{N} \sum\limits_{k=0}^{N-j} b_{jk} x_{mi}^{j} y_{mi}^{k} + \xi_i
\end{cases}, \quad i = 1, ..., n \tag{21}
$$

where n is the number of constructed correspondences, N is the order of polynomial, a and b are polynomial coefficients, (x_{si}, y_{si}) and (x_{mi}, y_{mi}) are the ith feature correspondence extracted from the slave and master images, and ς_i and ξ_i denote the normally distributed error terms with zero mean. Actually, (21) denote a 2D linear regression problem:

$$
\begin{cases}
x_{si} = \mathbf{X}_i^{\mathrm{T}} \boldsymbol{\theta} + \varsigma_i \\
y_{si} = \mathbf{X}_i^{\mathrm{T}} \boldsymbol{\psi} + \xi_i
\end{cases} \text{ with }
\begin{cases}
\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^{\mathrm{T}} = [a_{00}, a_{01}, \cdots, a_{N0}]^{\mathrm{T}} \\
\boldsymbol{\psi} = [\psi_1, \psi_2, \cdots, \psi_p]^{\mathrm{T}} = [b_{00}, b_{01}, \cdots, b_{N0}]^{\mathrm{T}} \\
\mathbf{X}_i = [X_{i1}, X_{i2}, \cdots, X_{ip}]^{\mathrm{T}} = [1, y_{mi}, \cdots, y_{mi}^{N}, x_{mi}, \cdots, x_{mi}^{N}]^{\mathrm{T}}
\end{cases}, \quad i = 1, ..., n \tag{22}
$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ are the unknown parameters to be estimated, and $p = (N + 1)(N + 2)/2$ denotes the number of unknowns. Then, the warp function estimation for SAR image registration can be transformed into the following optimization problems:

$$\begin{cases} Q_x := \text{Minimize} \sum_{i=1}^{h} (\mathbf{r_x^2})_i \\ Q_y := \text{Minimize} \sum_{i=1}^{h} (\mathbf{r_y^2})_i \end{cases} \text{with} \begin{cases} \mathbf{r_x} = [r_{x1}, r_{x2}, \cdots, r_{xn}]^T \\ \mathbf{r_y} = [r_{y1}, r_{y2}, \cdots, r_{yn}]^T \end{cases} \text{and} \begin{cases} r_{xi} = x_{si} - \mathbf{X}_i^T \boldsymbol{\theta} \\ r_{yi} = y_{si} - \mathbf{X}_i^T \boldsymbol{\psi} \end{cases}, \; i = 1, ..., n$$

$$(23)$$

where $(\mathbf{r_x^2})_i$ represents the ith element of the ordered squared residuals $(\mathbf{r_x^2})_1 \leq \cdots \leq (\mathbf{r_x^2})_i \leq \cdots \leq (\mathbf{r_x^2})_n$, and the meaning of $(\mathbf{r_y^2})_i$ can be likewise inferred. Each of the two optimizations in (23) is of the standard form (20). A direct solution to (23) may be thus achieved by decomposing 2D regression as two independent 1D regressions and using Fast-LTS to conduct estimation, respectively. This idea is feasible, but it may result in unnecessary computations because the feature positions in two image directions are in fact tied to each other, i.e., for the ith feature (x_i, y_i), the selection of x_i will naturally mean the selection of y_i. We can thus combine the two 1D regressions into a real 2D regression effectively, i.e., the extended Fast-LTS (EF-LTS):

Algorithm 3: EF-LTS

Step 31. Randomly draw p feature matches and LS fit them to estimate the initial parameters $\boldsymbol{\theta}_0$ and $\boldsymbol{\psi}_0$, and calculate the initial residuals $\mathbf{r_{0x}}$ and $\mathbf{r_{0y}}$ by

$$\begin{cases} \mathbf{r_{0x}} = [r_{0x1}, r_{0x2}, \cdots, r_{0xn}]^T \\ \mathbf{r_{0y}} = [r_{0y1}, r_{0y2}, \cdots, r_{0yn}]^T \end{cases} \text{and} \begin{cases} r_{0xi} = x_{si} - \mathbf{X}_i^T \boldsymbol{\theta}_0 \\ r_{0yi} = y_{si} - \mathbf{X}_i^T \boldsymbol{\psi}_0 \end{cases}, \; i = 1, ..., n. \qquad (24)$$

Then construct the initial h-subsets $\mathbf{Hx_0}$ and $\mathbf{Hy_0}$ by:

$$\begin{cases} \mathbf{Hx_0} = \{\boldsymbol{\pi}\mathbf{x}(1), \boldsymbol{\pi}\mathbf{x}(2), \cdots, \boldsymbol{\pi}\mathbf{x}(h)\} \subset \{1, 2, \cdots, n\} \\ \mathbf{Hy_0} = \{\boldsymbol{\pi}\mathbf{y}(1), \boldsymbol{\pi}\mathbf{y}(2), \cdots, \boldsymbol{\pi}\mathbf{y}(h)\} \subset \{1, 2, \cdots, n\} \end{cases} s.t. \begin{cases} (\mathbf{r_{0x}^2})_{\boldsymbol{\pi}\mathbf{x}(1)} \leq \cdots \leq (\mathbf{r_{0x}^2})_{\boldsymbol{\pi}\mathbf{x}(h)} \leq \cdots \leq (\mathbf{r_{0x}^2})_{\boldsymbol{\pi}\mathbf{x}(n)} \\ (\mathbf{r_{0y}^2})_{\boldsymbol{\pi}\mathbf{y}(1)} \leq \cdots \leq (\mathbf{r_{0y}^2})_{\boldsymbol{\pi}\mathbf{y}(h)} \leq \cdots \leq (\mathbf{r_{0y}^2})_{\boldsymbol{\pi}\mathbf{y}(n)} \end{cases}.$$

$$(25)$$

Carry out two C-steps on $\mathbf{Hx_0}$ and $\mathbf{Hy_0}$ to obtain the h-subsets $\mathbf{Hx_2}$ and $\mathbf{Hy_2}$ with smaller Q_x and Q_y, respectively. Iteratively repeat above procedures T times to obtain a set of h-subsets $\mathbf{Hx_2}$ and $\mathbf{Hy_2}$.

Step 32. Select 10 $\mathbf{Hx_2}$ with the smallest 10 Q_x and 10 $\mathbf{Hy_2}$ with the smallest 10 Q_y if T is larger than 10; otherwise, select all $\mathbf{Hx_2}$ and $\mathbf{Hy_2}$. Carry out C-steps on these h-subsets until convergence. The solutions corresponding to the smallest Q_x and Q_y are selected as the raw estimations $\boldsymbol{\theta}_\mathbf{r}$ and $\boldsymbol{\psi}_\mathbf{r}$, respectively.

Step 33. Calculate residuals $\mathbf{r_{rx}}$ and $\mathbf{r_{ry}}$ based on $\boldsymbol{\theta}_\mathbf{r}$ and $\boldsymbol{\psi}_\mathbf{r}$,

$$\begin{cases} \mathbf{r_{rx}} = [r_{rx1}, r_{rx2}, \cdots, r_{rxn}]^T \\ \mathbf{r_{ry}} = [r_{ry1}, r_{ry2}, \cdots, r_{ryn}]^T \end{cases} \text{with} \begin{cases} r_{rxi} = x_{si} - \mathbf{X}_i^T \boldsymbol{\theta}_\mathbf{r} \\ r_{ryi} = y_{si} - \mathbf{X}_i^T \boldsymbol{\psi}_\mathbf{r} \end{cases}, \; i = 1, ..., n \qquad (26)$$

and estimate the error scales σ_x and σ_y by

$$\begin{cases} \sigma_x = C_1 \sqrt{\dfrac{1}{h} \sum_{i=1}^{h} (\mathbf{r_{rx}^2})_i} \\ \sigma_y = C_2 \sqrt{\dfrac{1}{h} \sum_{i=1}^{h} (\mathbf{r_{ry}^2})_i} \end{cases} \qquad (27)$$

where C_1 and C_2 are correction factors to achieve consistency at Gaussian error distributions [50]. Based on (27), we further calculate two weights by:

$$w_{xi} = \begin{cases} 1 & \text{if } |r_{xi}/\sigma_x| \le 2.5 \\ 0 & \text{if } |r_{xi}/\sigma_x| > 2.5 \end{cases} \quad w_{yi} = \begin{cases} 1 & \text{if } |r_{yi}/\sigma_y| \le 2.5 \\ 0 & \text{if } |r_{yi}/\sigma_y| > 2.5 \end{cases}, \quad i = 1, ..., n. \quad (28)$$

The credible correspondence in both directions of x and y is chosen:

$$w_i = w_{xi} \,\&\, w_{yi}, \quad i = 1, ..., n \quad (29)$$

where "&" denotes the logical AND operator. The final estimations $\boldsymbol{\theta_f}$ and $\boldsymbol{\psi_f}$ are attained by LS solving the following optimizations:

$$\begin{cases} \boldsymbol{\theta_f} = \text{argmin} \sum_{i=1}^{n} w_i r_{xi}^2 \\ \boldsymbol{\psi_f} = \text{argmin} \sum_{i=1}^{n} w_i r_{yi}^2 \end{cases} \quad (30)$$

which in fact indicates the weighted LS.

Step 33 makes EF-LTS obtain more accurate and stable estimation than the original LTS. The logical AND in (29) shows that only the feature correspondence which is correctly matched in both x- and y-direction is considered as an inlier. This is necessary for accurate estimation because mismatching in one direction may also affect the matching in another. The bound in (28) is set as 2.5 for there are very few residuals larger than 2.5σ in a Gaussian situation [50].

In Fast-LTS, the random sampling number T is a constant 500. This is inappropriate because accurate estimation only requires one p-match to being "clean." Let q denote the percentage of inliers in data, then the probability ε of having at least one "clean" p-match among all the T random p-matches can be expressed as

$$\varepsilon = 1 - (1 - q^p)^T. \quad (31)$$

Since the trimming constant h is chosen beforehand according to the percentage of inliers, a good estimation of q can be obtained by

$$\hat{q} = \frac{h}{n}. \quad (32)$$

Therefore, if a required false alarm rate ε for the estimation is given, the sampling number T can be then calculated by combining (31) and (32):

$$T = \left\lceil \frac{\log(1 - \varepsilon)}{\log\left(1 - \left(\frac{h}{n}\right)^p\right)} \right\rceil. \quad (33)$$

Thus, iteration in EF-LTS is controlled by the inlier percentage rather than the inlier number. **Table 4** shows the sampling number T under given N and q when $\varepsilon = 0.99$. It can be seen that even the worst sampling number 293 is much smaller than 500 for $N = 2$. Thus, the constant 500 sampling will be redundant for the second-order polynomial, but will be insufficient for the third-order polynomial with smaller q, as listed in **Table 4**.

The inlier percentage q is in fact related to *MFAR* by:

$$q = 1 - MFAR. \quad (34)$$

Thus, besides introducing more iterations and computation load, higher *MFAR* will also lead to a smaller h-subset, which indicates more localization and less

N	q						
	0.5	0.6	0.7	0.75	0.8	0.9	0.95
0	7	6	4	4	3	2	2
1	35	19	11	9	7	4	3
2	293	97	37	24	16	7	4
3	4714	760	161	80	41	11	6

Table 4.
Sampling number T under different inlier percentage q and polynomial order N when ε = 0.99.

accuracy in estimation and worse consistency between the extracted features and retrieval parameters. This is why *FH*-15(-2) can achieve better *ATE* than *FH*-9(-1), as displayed in **Table 2**. As presented in Section 3, on *MFAR* and many other criteria, SURF is identified to be the best for general SAR image registration. SURF may thus also improve the efficiency and accuracy of parameter retrieval besides the good performance on feature extraction and matching.

When the correspondence number n is large, a similar nested extension can be also taken for EF-LTS by randomly partitioning the correspondences into M subsets with equal cardinality, and the trimming constant h_s and sampling number T_s of each subset should be also reduced by M times relative to h and T. On each subset, we first implement Step 31 for T_s h_s-subsets of $\mathbf{Hx_2}$ and $\mathbf{Hy_2}$. Based on which we then implement Step 32 and Step 33 on all the constructed correspondences with original h and T. In this way, an efficient retrieval can be still achieved.

To evaluate EF-LTS for SAR image registration, we also use it to the InSAR image pair given in **Figure 4(a)** and **(b)**. Similarly, the feature correspondences are first constructed by SURF with *HF*-9(-1), then we run EF-LTS 100 times to retrieve the affine parameters and calculate CC and SNR. The obtained parameters, CC, and SNR of each execution are shown in **Figure 5**, while the mean and standard deviation of the parameters, CC, and SNR are listed in **Table 3**. It is revealed that EF-LTS behaves very stable and the estimated parameters, CC, and SNR are invariant for each execution. It can reach an averagely better CC and SNR than RANSAC and is more appropriate for InSAR image registration. **Figure 4(c)** and **(d)** further illustrates the interferogram and correlation map of the coregistered InSAR pair with wrap parameters estimated by EF-LTS. Interferogram is the argument or phase of the dot production between the complex master image and the complex conjugation of the registered slave image, while correlation map measures CC of the 3×3 patches around each corresponding pixel position between the images. The interferogram fringe is clear and the correlation is strong in stable area such as the brighter buildings in **Figure 4(a)** and **(b)** and the upper-right bare land. But in the upper-left residential area, the interferogram becomes less clear and the correlation is relatively small probably because the scattering is very sensitive to incidence changes. While in other area (mainly vegetable lands and parking lot), the interferogram is almost lost and the coherence is very low due to the temporal and/or volume decorrelation. All these match with the ground truth very well.

5. Experiment and analysis

Based on the finding in Sections 2–4, we propose to conduct high accurate SAR image registration by using EF-LTS to fit the SURF correspondences. The scheme works as follows:

Algorithm 4: Accurate SAR image registration based on SURF features and EF-LTS

Step 41. Use *FH*-9(-*Fs*) to extract SURF keypoints from master and slave images, respectively.

Step 42. Construct initial feature correspondences by simply matching SURF descriptors.

Step 43. Robustly processing the correspondences with EF-LTS to retrieve the warp function.

Step 44. Transform and interpolate the slave image to geometrically align it to master image.

Actually, this scheme has been put into practice in the above experiments. In this section, we further devise an experiment to check it on MiniSAR pair. The images we use are two high-resolution SAR images of the entrance gate of the Sandia Research Park acquired by the Ku-Band MiniSAR system developed by the Sandia Laboratory [51]. The images are taken from different tracks with different incidences and squints, as listed in **Table 5**, while the platform altitude is just beyond 1 km. All these reveal the nontrivial target relief-induced geometrical warping between images, which, however, cannot be compensated beforehand for lack of ground truth such as DEM and target height. Besides this, the images also experience a very large intensity variation. To enhance the texture, we use the logarithmic intensity of original complex images, as shown in **Figure 7(a)** and **(b)**. To achieve a more precise approximation to the real warping, we divide the image pair into four 500 × 500 patch pairs. The geometrical warping on each patch pair is approximated as an affine transformation (the higher order polynomial has also been used to model the warp function, but unsatisfactory registration result is attained). We adopt *HF*-9(-4) SURF detector to extract feature correspondences from each patch pair, and EF-LTS is then used to obtain the affine parameters, based on which the slave image is finally aligned to the master image. To illustrate the registration accuracy, we fuse and overlap the coregistered images together. The RGB fusion in **Figure 7(c)** is obtained by treating the master image and the coregistered slave image as red and green, respectively, while zeroing the blue component. The well-distributed yellow then immediately illustrates the accurate registration of the images. The overlapping in **Figure 7(d)** is obtained by simply averaging the two coregistered images. It contains the whole information of the two images but has fewer speckles.

To further evaluate the registration performance of the scheme, in the following we focus on the two pole-like target areas 1 and 2 in **Figure 7(d)** with their corresponding Google optical images shown in **Figure 8(g)** and **(h)**, respectively. **Figure 8(i)** portrays the details of Pole 2 in the Street View of Google Maps. The

Parameters	Master image	Slave image
Azimuth resolution	0.1016 m	0.1016 m
Range resolution	0.1016 m	0.1016 m
Grazing angle	27.0107°	26.1892°
Global track angle	158.3687°	153.0825°
Central frequency	16.8 GHz	16.8 GHz
Platform altitude	1.6715 km	1.6715 km
Squint	−89.9935°	−89.9924°

Table 5.
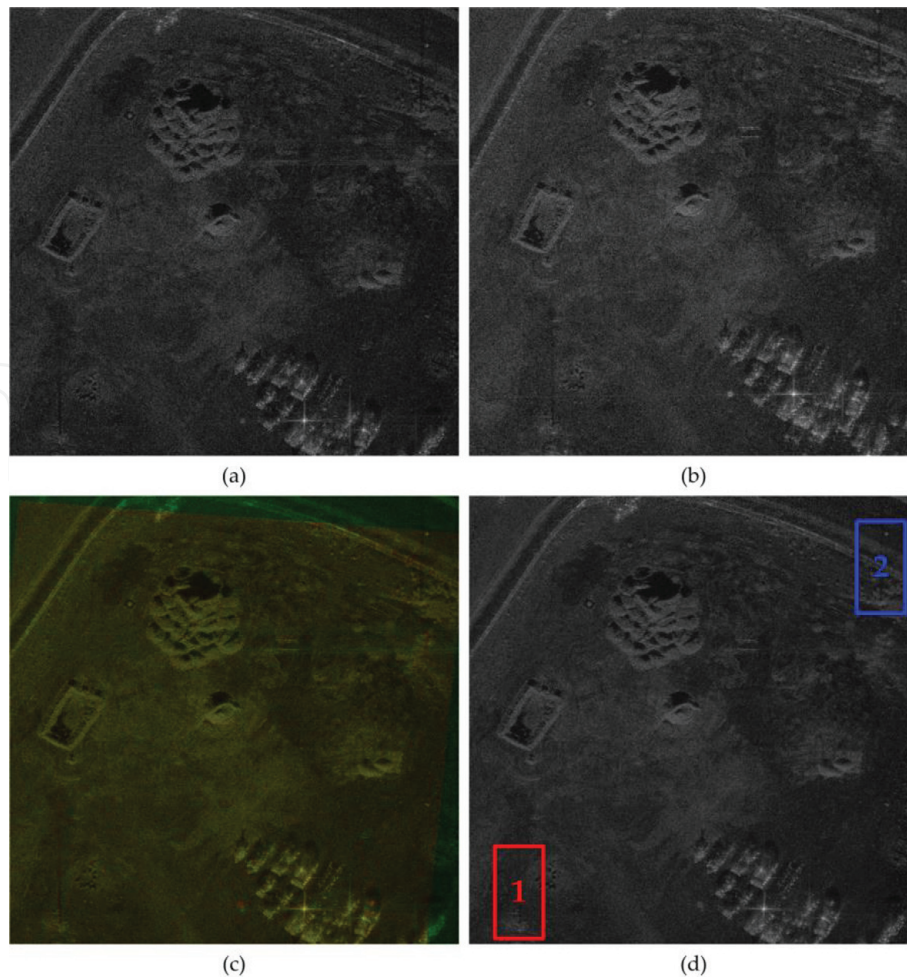*Imaging parameters of the two MiniSAR images.*

Figure 7.
Registration of the MiniSAR image pair. (a) Master image, (b) slave image, and (c) pseudocolor fusion as well as (d) the overlapping of them after registration with EF-LTS to fit the SURF features. "1" and "2" in (d) indicate two pole-like targets which are further detailed in **Figure 8**.

target is shown to be the power transmission pole. **Figure 8(a)–(c)** exhibits the SAR imagery of Pole 1 in the master image, coregistered slave image, and overlapped image, respectively. The corresponding SAR imageries of Pole 2 are displayed in **Figure 8(d)–(f)**, respectively. It is known that the darker pole-like feature in each SAR image is not the real pole scattering, but its shadow under the irradiation of radar. The actual scattering center of the pole is overlapped with its ground position because of the dominant dihedral backscattering between the pole and ground. From **Figure 8(c)** and **(f)**, we can find that the shadows of the two poles are still separated after registration due to the volume-induced warping. According to our estimate, the separations are about 6.5 and 5°, respectively, which approach to the actual track angle 5.2862°. Except for these shadows, the poles and other area are accurately overlapped. Nice registration is still achieved despite the large local distortion and decorrelation. Moreover, the experiment also validates the strategy for general feature-based SAR image registration, i.e., to focus on the global regis-tration and to neglect the local discontentment. The accurate registration of each pixel is impossible and unnecessary. It should be noted that the conventional SAR image registrations including the feature-based approaches focused in current chapter are mainly appropriate for images with approximated low-order polyno-mial geometrical warping. For SAR images taken from area of rough topography with long baseline, we need some more complex approaches with the *a priori* ground truth information being included, such as the DEM-assisted registration [48]. Although the SAR and InSAR image pairs used in the experiment are all
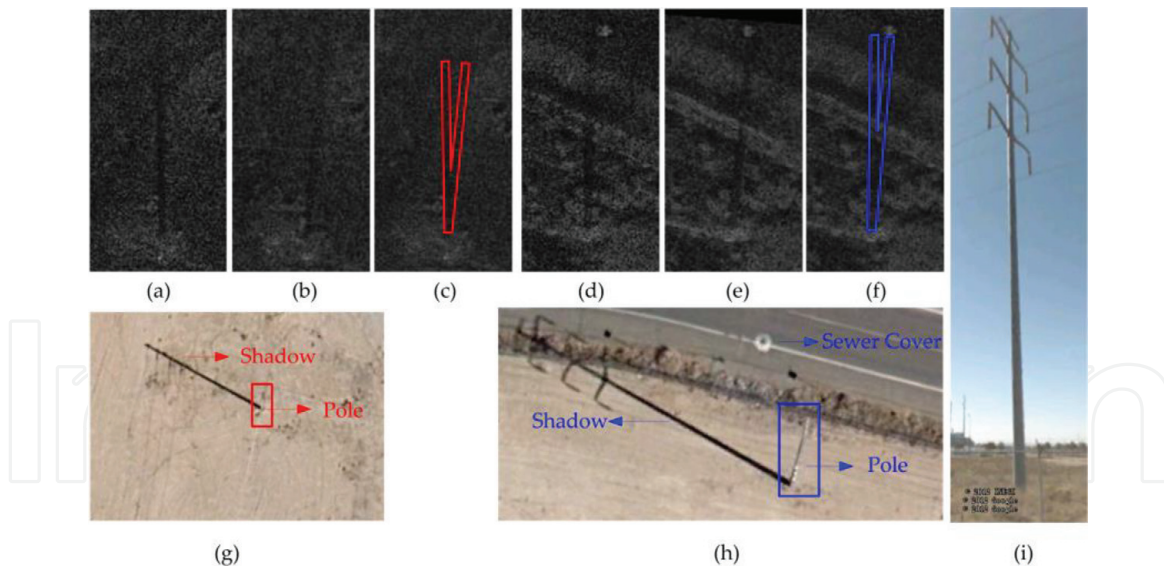
Figure 8.
Registration of the two pole-like targets. SAR imagery of pole "1" in (a) master image, (b) coregistered slave image, and (c) overlapped image, as well as (g) the corresponding Google optical image. SAR imagery of pole "2" in (d) master image, (e) coregistered slave image, and (f) overlapped image, as well as (h) its Google optical image and (i) detailed portrayal in Google Maps.

monopolarized, the developed scheme is also appropriate to the registration of fully polarimetric SAR (PolSAR) images. Different from monopolarized SAR, each cell in PolSAR image is a scattering matrix $\mathbf{S}$ with four entries S_{HH}, S_{HV}, S_{VH}, and S_{VV} [52]:

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix}. \tag{35}$$

Nevertheless, by taking the squared Frobenius norm of matrix $\mathbf{S}$ [53]:

$$SPAN = \|\mathbf{S}\|_F^2 = |S_{HH}|^2 + |S_{HV}|^2 + |S_{VH}|^2 + |S_{VV}|^2 \tag{36}$$

we can then obtain the total power (also known as *SPAN*) of target. An accurate registration of PolSAR images can be eventually achieved by simply using the developed scheme to the corresponding *SPAN* image pair.

6. Conclusion

SAR coherent imaging unavoidably brings about geometrical distortion and speckle into the acquired images and makes the registration of SAR images much more complicated. In this chapter, we focus on two important procedures in general feature-based SAR registration, i.e., the feature extraction and the parameter retrieval by identifying the appropriate feature and the appropriate estimation algorithm. As for the former, we conduct a detailed evaluation on the commonly used features such as tie points, Harris corner, SIFT, and SURF. We find that SURF outperforms others in terms of the geometrical invariance of feature, extraction speed, accuracy of localization, geometrical invariance of descriptor, matching speed, robustness to decorrelation, and flexibility to image speckling. Among these criteria, feature's flexibility to speckle is particularly focused because speckle impacts the feature extraction and matching, while speckle filtering may change the feature position and impact the subpixel localization. The Fast-Hessian detector of

SURF has a potential relation with the refined Lee speckle filter. SSP in SURF just indicates that we use a series of box filters of different size to filter speckles and extract features of different scales. Thus, SURF is very flexible to deal with SAR speckle. In view of the application with strict requirement for registration accuracy, we suggest using the SURF detector of *HF-9(-1)* to the *Fs* times interpolated images with unchanged sampling step to extract feature. The new detector *HF-9(-Fs)* can significantly improve the registration accuracy to subpixel (<1 pixel) and is especially fit for high accurate SAR image registration.

Parameter retrieval in SAR registration is difficult because spatial or temporal decorrelation will always introduce mismatches into the obtained feature correspondences. The estimator should be robust to outliers. We find that the commonly used RANSAC may trap into local occlusion and result in uncertain parameter retrieval. This uncertainty is more severe in SAR image registration because SAR geometrical warping varies from pixel to pixel, but the low-order polynomial approximation can only account for global registration instead of the local contentment. The local-to-global strategy in RANSAC may thus magnify the local distortion, aggravate the estimation uncertainty, and damnify the global registration accuracy although a largest CS is obtained. To achieve a stable registration for SAR images, we should estimate the parameters with more correspondences to reflect the true support than just a MSS, and apply an appropriate loss function. This leads us to EF-LTS, which improves Fast-LTS from 1D regression to 2D regression, and provides us an adaptive determination of the number of random sampling instead of setting it as a constant 500. EF-LTS conducts registration by LS fitting at least half of the correspondences to minimize the squared residual. It behaves very stable and is averagely better than RANSAC. Hence, we recommend conducting SAR image registration by fitting SURF features with EF-LTS. Experiments on both InSAR and MiniSAR image pairs validate the nice performance of this registration scheme.

Acknowledgements

Author details

Dong Li[1], Yunhua Zhang[1,2]* and Xiaojin Shi[1]

1 Key Laboratory of Microwave Remote Sensing, National Space Science Center, Chinese Academy of Sciences, Beijing, China

2 University of Chinese Academy of Sciences, Beijing, China

*Address all correspondence to: zhangyunhua@mirslab.cn

IntechOpen

References

[1] Homer J, Longstaff ID. Minimising the tie patch window size for SAR image coregistration. Electronics Letters. 2003;**39**:122-124. DOI: 10.1049/el:20030054

[2] Zou W, Li Y, Li Z, Ding X. Improvement of the accuracy of InSAR image co-registration based on tie points: A review. Sensors. 2009;**9**: 1259-1281. DOI: 10.3390/s90201259

[3] Giles AB, Massom RA, Warner RC. A method for sub-pixel scale feature-tracking using Radarsat images applied to the Mertz Glacier Tongue, East Antarctica. Remote Sensing of Environment. 2009;**113**:1691-1699. DOI: 10.1016/j.rse.2009.03.015

[4] Li D, Zhang Y. A fast offset estimation approach for InSAR image subpixel registration. IEEE Geoscience and Remote Sensing Letters. 2012;**9**: 267-271. DOI: 10.1109/LGRS.2011. 2166752

[5] Harris C, Stephens M. A combined corner and edge detector. In: Proceedings of the Alvey Vision Conference (AVC'88); 31 August–2 September 1998; Manchester, U.K. pp. 147-151

[6] Bentoutou Y, Taleb N, Kpalma K, Ronsin J. An automatic image registration for applications in remote sensing. IEEE Transactions on Geoscience and Remote Sensing. 2005; **43**:2127-2137. DOI: 10.1109/TGRS. 2005.853187

[7] Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision. 2004;**60**:91-110. DOI: 10.1023/B: VISI.0000029664.99615.94

[8] Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF).

Computer Vision and Image Understanding. 2008;**110**:346-359. DOI: 10.1016/j.cviu.2007.09.014

[9] Lindeberg T. Feature detection with automatic scale selection. Internatioal Journal of Computer Vision. 1998;**30**: 79-116. DOI: 10.1023/A:1008045108935

[10] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2005;**27**:1615-1630. DOI: 10. 1109/TPAMI.2005.188

[11] Wessel B, Huber M, Roth A. Automatic image-to-image registration of near real-time SAR images. In: Proceedings of the Envisat Symposium; 23–27 April 2007; Montreux, Switzerland. pp. 1-6

[12] Li F, Zhang G, Yan J. Coregistration based on SIFT algorithm for synthetic aperture radar interferometry. In: Archives of International Society got Photogrammetry and Remote Sensing (ISPRS'08); 3–11 July 2008; Beijing, China. pp. 123-128

[13] Chen E, Li Z, Tian X, Li S. Application of scale invariant feature transformation to SAR imagery registration. Acta Automatica Sinica. 2008;**34**:861-868. DOI: 10.3724/SP. J.1004. 2008.00861

[14] Li D, Zhang Y. Geometric feature-based image coregistration approach for InSAR. In: Proceedings of the Asian-Pacific Conference on Synthetic Aperture Radar; 26–30 October, 2009; Xi'an, China. pp. 1026-1030

[15] Schwind P, Suri S, Reinartz P, Siebert A. Applicability of the SIFT operator to geometric SAR image registration. International Journal of

Remote Sensing. 2010;**31**:1959-1980. DOI: 10.1080/01431160902927622

[16] Wang S, You H, Fu K. BFSIFT: A novel method to find feature matches for SAR image registration. IEEE Geoscience and Remote Sensing Letters. 2012;**9**:649-653. DOI: 10.1109 / LGRS.2011.2177437

[17] Gong M, Zhao S, Jiao L, Tian D, Wang S. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual Information. IEEE Transactions on Geoscience and Remote Sensing. 2014;**52**:4328-4338. DOI: 10.1109/TGRS.2013.2281391

[18] Dellinger F, Delon J, Gousseau Y, Michel J, Tupin F. SAR-SIFT: A SIFT-like algorithm for SAR images. IEEE Transactions on Geoscience and Remote Sensing. 2015;**53**:453-466. DOI: 10.1109/TGRS.2014.2323552

[19] Wang F, You H, Fu X. Adapted anisotropic Gaussian SIFT matching strategy for SAR registration. IEEE Geoscience and Remote Sensing Letters. 2015;**12**:160-164. DOI: 10. 1109/LGRS.2014.2330593

[20] Wang B, Zhang J, Lu L, Huang G, Zhao Z. A uniform SIFT-like algorithm for SAR image registration. IEEE Geoscience and Remote Sensing Letters. 2015;**12**:1426-1430. DOI: 10.1109/LGRS.2015.2406336

[21] Zeng L, Zhou D, Liang J, Zhang K. Polar scale-invariant feature transform for synthetic aperture radar image registration. IEEE Geoscience and Remote Sensing Letters. 2017;**14**:1101-1105. DOI: 10.1109/LGRS.2017.2698450

[22] Ma W, Wen Z, Wu Y, Jiao L, Gong M, Zheng Y, et al. Remote sensing image registration with modified SIFT and enhanced feature matching. IEEE Geoscience and Remote Sensing Letters. 2017;**14**:3-7. DOI: 10.1109/LGRS.2016.2600858

[23] Xiang Y, Wang F, You H. OS-SIFT: A robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas. IEEE Transactions on Geoscience and Remote Sensing. 2018;**56**:3078-3090. DOI: 10.1109/TGRS.2018.2790483

[24] Teke M, Temizel A. Multi-spectral satellite image registration using scale-restricted SURF. In: Proceedings of International Conference on Pattern Recognition (ICPR'10); 23–26 August 2010; Istanbul, Turkey. pp. 2310-2313

[25] Nguyen H-G, Fablet R, Ehrhold A, Boucher J-M. Keypoint-based analysis of sonar images: Applications to seabed recognition. IEEE Transactions on Geoscience and Remote Sensing. 2012;**50**:1171-1184. DOI: 10.1109/TGRS.2011.2165848

[26] Liu R, Wang Y. SAR image matching based on speeded up robust feature. In: Proceedings of the WRI Global Congress on Intelligent System (GCIS'09); 19–21 May, 2009; Xiamen, China. pp. 518-522

[27] Lee WK, Kim AL. An efficient automatic geo-registration technique for high resolution spaceborne SAR image fusion. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGASS'11); 24–29 July, 2011; Vancouver, Canada. pp. 3566-3569

[28] Zhang H, Ni W, Yan W, Wu J, Li S. Robust SAR image registration based on edge matching and refined coherent point drift. IEEE Geoscience and Remote Sensing Letters. 2015;**12**:2115-2119. DOI: 10.1109/LGRS.2015.2451396

[29] Li D, Zhang Y. On the appropriate feature for general SAR image registration. Proceedings of SPIE. 2012; **8536**:85360X. DOI: 10.1117/12.970520

[30] Fischler MA, Bolles RC. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM. 1981;**24**: 381-395. DOI: 10.1145/358669.358692

[31] Zuliani M. Computational methods for automatic image registration [thesis]. Santa Barbara, U.S.: University of California at Santa Barbara; 2007

[32] Massart DL, Kaufman L, Rousseeuw PJ, Leroy A. Least median of squares: A robust method for outlier and model error detection in regression and calibration. Analytica Chimica Acta. 1986;**187**:171-179. DOI: 10.1016/S0003-2670(00)82910-4

[33] Rousseeuw PJ, Van Driessen K. Computing LTS regression for large data sets. Data Mining and Knowledge Discovery. 2006;**12**:29-45. DOI: 10.1007/s10618-005-0024-4

[34] Zhang Z, Deriche R, Faugeras O, Luong QT. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence. 1995;**78**:87-119. DOI: 10.1016/0004-3702(95)00022-4

[35] Li D, Zhang Y. The appropriate parameter retrieval algorithm for feature-based SAR image registration. Proceedings of SPIE. 2012;**8536**:85360Y. DOI: 10.1117/12.970522

[36] Brown LG. A survey of image registration techniques. ACM Computing Surveys. 1992;**24**:325-376. DOI: 10.1145/146370.146374

[37] Zitova B, Flusser J. Image registration methods: A survey. Image and Vision Computing. 2003;**21**:

977-1000. DOI: 10.1016/S0262-8856(03)00137-9

[38] Dawn S, Saxena V, Sharma B. Remote sensing image registration techniques: A survey. In: Proceedings of International Conference on Image and Signal Processing (ICISP'10); 30 June–2 July, 2010; Trois-Rivieres, Canada. pp. 103-112

[39] Mikolajczyk K, Schmid C. An affine invariant interest point detector. In: Proceedings of European Conference on Computer Vision (ECCV'02); 28–31 May, 2002; Copenhagen, Denmark. pp. 128-142

[40] Rufino G, Moccia A, Esposito S. DEM generation by means of ERS tandem data. IEEE Transactions on Geoscience and Remote Sensing. 1998; **36**:1905-1912. DOI: 10.1109/36. 729362

[41] Li FK, Goldstein RM. Studies of multibaseline spaceborne interferometric synthetic aperture radars. IEEE Transactions on Geoscience and Remote Sensing. 1990; **28**:88-97. DOI: 10.1109/36.45749

[42] Brown M, Lowe DG. Invariant features from interest point groups. In: Proceedings of the British Machine Vision Conference (BMVC'02); 2–5 September, 2002; Cardiff, U.K. pp. 253-262

[43] Li D, Zhang Y. A novel approach for the registration of weak affine images. Pattern Recognition Letters. 2012;**33**: 1647-1655. DOI: 10.1016/j. patrec.2012.04.009

[44] Li Z, Bethel J. Image coregistration in SAR interferometry. In: Archives of International Society got Photogrammetry and Remote Sensing (ISPRS'08); 3–11 July 2008; Beijing, China. pp. 433-438

[45] Bamler R, Eineder M. Accuracy of differential shift estimation by

correlation and split-bandwidth interferometry for wideband and Delta-k SAR systems. IEEE Geoscience and Remote Sensing Letters. 2005;**2**:151-155. DOI: 10.1109/LGRS.2004.843203

[46] Lee JS, Pottier E. Polarimetric Radar Imaging: From Basics to Applications. Boca Raton, U.S.: CRC Press; 2009. pp. 143-177

[47] Li D, Zhang Y. A rigorous SAR epipolar geometry modeling and application to 3D target reconstruction. IEEE Selected Topics in Applied Earth Observations and Remote Sensing. 2013; **6**:2316-2323. DOI: 10.1109/ JSTARS.2013.2249575

[48] Nitti DO, Hanssen RF, Refice A, Bovenga F, Nutricato R. Impact of DEM-assisted coregistration on high-resolution SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing. 2011;**49**:1127-1143. DOI: 10.1109/TGRS.2010.2074204

[49] Gabriel AK, Goldstein RM. Crossed orbit interferometry: Theory and experimental results from SIR-B. International Journal of Remote Sensing. 1988;**9**:857-872. DOI: 10. 1080/ 01431168808954901

[50] Rousseeuw PJ, Leroy A. Robust Regression and Outlier Detection. Hoboken, U.S.: John Wiley & Sons, Inc.; 2005. pp. 1-157. DOI: 10.1002/ 0471725382

[51] Sweet AD, Dubbert DF, Doerry AW, Sloan GR, Dee Gutierrez V. A portfolio of fine resolution Ku-band MiniSAR images. In: Proceedings of SPIE Defense and Security Symposium; 17–21 April, 2006; Orlando, U.S. pp. 621002-621006

[52] Li D, Zhang Y. Unified Huynen phenomenological decomposition of radar targets and its classification applications. IEEE Transactions on Geoscience and Remote Sensing. 2016;

54:723-743. DOI: 10.1109/TGRS. 2015.2464113

[53] Li D, Zhang Y. Adaptive model-based classification of PolSAR data. IEEE Transactions on Geoscience and Remote Sensing. 2018;**56**:1-16. DOI: 10.1109/TGRS.2018.2845944