


An Advanced Rotation Invariant Descriptor for SAR Image Registration

Yuming Xiang ^{1,2,*} , Feng Wang ¹, Ling Wan ^{1,2} and Hongjian You ^{1,2}

¹ Key Laboratory of Technology in Geo-spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China; wfeng_gucas@126.com (F.W.); wanling15@mails.ucas.ac.cn (L.W.); hjyou@mail.ie.ac.cn (H.Y.)

² University of Chinese Academy of Sciences, Beijing 100000, China

* Correspondence: z199208081010@163.com; Tel.: +86-152-1056-0903

Received: 25 May 2017; Accepted: 2 July 2017; Published: 4 July 2017

Abstract: The Scale-Invariant Feature Transform (SIFT) algorithm and its many variants have been widely used in Synthetic Aperture Radar (SAR) image registration. The SIFT-like algorithms maintain rotation invariance by assigning a dominant orientation for each keypoint, while the calculation of dominant orientation is not robust due to the effect of speckle noise in SAR imagery. In this paper, we propose an advanced local descriptor for SAR image registration to achieve rotation invariance without assigning a dominant orientation. Based on the improved intensity orders, we first divide a circular neighborhood into several sub-regions. Second, rotation-invariant ratio orientation histograms of each sub-region are proposed by accumulating the ratio values of different directions in a rotation-invariant coordinate system. The proposed descriptor is composed of the concatenation of the histograms of each sub-region. In order to increase the distinctiveness of the proposed descriptor, multiple image neighborhoods are aggregated. Experimental results on several satellite SAR images have shown an improvement in the matching performance over other state-of-the-art algorithms.

Keywords: SAR; image registration; feature descriptor; rotation invariance

1. Introduction

Synthetic Aperture Radar (SAR) image registration is the fundamental task of many image applications, such as image fusion, image mosaicking, change detection, and so on. Because of the extensive use of SAR images, SAR image registration becomes increasingly important. The registration algorithms can be roughly divided into two categories: intensity-based and feature-based. Affected by speckle noise and different imaging conditions, the intensity and geometric information of the same ground scene in SAR images differ widely. Consequently, feature-based methods with some particular invariance may be more suitable than intensity-based ones for SAR image registration [1].

Most of the feature-based methods consist of three steps: keypoints' detection, keypoints' matching and transformation model estimation. First, feature-based methods detect significant points that correspond to distinctive points of the same scene in two images, such as corner points, line intersections and centroid pixels of close-boundary regions [2,3]. Second, each feature point from one image (called the reference image) is matched with the corresponding point of the other image (called the sensed image) by various feature descriptors or similarity measures along with spatial relationships among the keypoints, such as the famous Scale-Invariant Feature Transform (SIFT) descriptor [4], shape context [5] and spectral graph [6]. Third, due to the complex nature of SAR images, the matched keypoints often result in a high number of false matches, which have a significant impact on determining the transformational model [7]. Therefore, robust algorithms are

applied to remove outliers, such as RANdom SAmple Consensus (RANSAC) [8] and A Contrario RANdom SAmple Consensus (AC-RANSAC) [9]. Then, a transformation model is estimated using the correctly-matched keypoints.

Among the feature-based methods, SIFT-like algorithms are the most widely-used techniques due to the efficient performance and invariance to scale, rotation and illumination changes. However, the traditional SIFT algorithm does not perform well on the SAR images due to the effect of speckle noise [1]. Several improvements have been proposed to improve the SIFT algorithm for SAR image registration. Some algorithms ameliorated the SIFT algorithm by extracting features starting from the second octave [10], skipping the dominant orientation assignment when the matching images do not have rotation transformation [11], or replacing the Gaussian filter with several anisotropic filters [12], or designing a new gradient specifically dedicated to SAR images by utilizing the Ratio Of the Exponentially-Weighted Average (ROEWA) instead of a differential (SAR-SIFT) [1]. In SIFT-like algorithms, rotation invariance is achieved by assigning a dominant orientation to each keypoint. However, Fan et al. [13] pointed out that the computed orientation is not stable enough and adversely affects the matching performance of the SIFT descriptor. Moreover, the calculation of the dominant orientation in SAR images is strongly affected by speckle noise.

In order to solve the aforementioned problems, we propose a robust feature descriptor for SAR image registration, which combine the advantages of the ratio-based detectors [14] and the intensity order pooling [13]. Considering the inherent property of SAR images, an improved intensity order pooling method is introduced to partition the circular neighborhood, then the rotation-invariant sub-regions of image neighborhood are obtained. For each sub-region, we propose a rotation-invariant ratio orientation histogram, which is obtained by accumulating the ratio values of different directions in a rotation-invariant coordinate system. The proposed descriptor is composed of the concatenation of the histograms of each sub-region.

The main contributions of the paper are given as follows: (1) a rotation-invariant local descriptor is constructed without assigning a dominant orientation, which can be widely used in feature-based registration techniques; (2) considering the inherent property of SAR images, we proposed a rotation-invariant ratio orientation histogram that is robust to speckle noise. Additionally, the intensity order pooling method is improved to adapt to SAR images. Rotation invariance is achieved by the improved intensity orders and rotation-invariant ratio orientation histograms; and (3) multiple image neighborhoods are aggregated to increase the distinctiveness of the descriptor.

2. Methodology

2.1. Local Descriptors for SAR Image Registration

Local descriptors have achieved good performance for optical image registration, such as SIFT [4], Speeded Up Robust Features (SURF) [15], Oriented FAST and Rotated BRIEF (ORB) [16], shape-based invariant texture feature [17] and so on. However, when these descriptors are directly applied to SAR images, the results are poor due to the effect of speckle noise and large differences between the intensity and geometric information [10]. Consequently, we only focus on the local descriptors that have been successfully operated on SAR images. As mentioned before, some improved SIFT-like descriptors have shown good performance on SAR image registration. Among them, SAR-SIFT is the state-of-the-art method. Benefiting from the Constant False Alarm Rate (CFAR) property of ratio-based edge detectors [14], the gradient by ratio used in SAR-SIFT is robust to speckle noise, leading to a good performance specifically on SAR images. The SAR-SIFT descriptor utilizes the gradient by ratio to compute gradient orientation for each point. Then, dominant orientation is assigned to each keypoint to maintain the rotation invariance. More details about SAR-SIFT can be seen in [1].

Other feature descriptors have also been developed. Dai et al. [18] found matches between SAR and optical images using improved chain-code representation and invariant moments. Yasein et al. [19]

obtained correspondences between the feature points using Zernike moments. Wong et al. [20] proposed an algorithm that makes use of phase-congruency moment-based patches as local-feature descriptors. Huang et al. [5] improved the shape context descriptor to make it fit for use with complex remote sensing images. Recently, learning-based feature descriptors are applied to SAR image processing. Yang et al. [21] introduced a computationally efficient graphical model for densely labeling large remote sensing images, which combined the advantages of the multiscale visual features and hierarchical smoothing. Hu et al. [22] generated image features via extracting CNN features from different layers. Yang et al. [23] proposed a high-level feature learning method to describe an image sample. Hu et al. [24] presented an improved unsupervised feature learning algorithm based on spectral clustering, which can not only adaptively learn good local feature representations but also discover intrinsic structures of local image patches. The aforementioned local descriptors almost all adapted a dominant orientation to achieve rotation invariance. However, Fan et al. [13] and Liu et al. [25] claimed that the dominant orientation assignment based on local image statistics is an error-prone process and it will make many true corresponding points un-matchable by their descriptors. Moreover, the calculation of dominant orientation is not stable enough due to the speckle noise of SAR images.

In order to assess this, an orientation error experiment is operated on 40 pairs of SAR images with rotation differences. We choose the SAR-SIFT descriptor as representative. The Orientation Error (OE) for two corresponding points is given as: $\xi_{ori} = (\theta' - \theta) - \theta(H)$, where θ is the dominant orientation of the keypoint $(x, y; \theta)$ in the reference image and θ' is the dominant orientation of the corresponding keypoint $(x', y'; \theta')$ in the sensed image. $\theta(H)$ is the orientation difference between two images according to the ground truth model H . Figure 1a shows the distribution of OEs between all corresponding points. Figure 1b shows the distribution of OEs between corresponding points that are correctly matched. We can see that the OEs of the correctly-matched points are almost in the range of $[-20^\circ, 20^\circ]$. However, the rate of corresponding points that have the OEs in the range of $[-20^\circ, 20^\circ]$ is 62.59%. This means that 37.41% of the corresponding points may not be correctly matched. Consequently, the matching performance can be significantly improved by a more accurate estimation of dominant orientation or building a rotation-invariant local descriptor without assigning a dominant orientation.

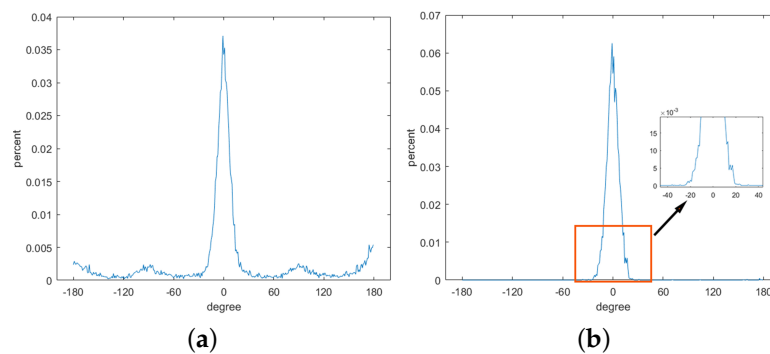


Figure 1. (a) orientation error (OE) distribution of all of the corresponding points; (b) OE distribution of the correctly-matched points.

2.2. The Proposed Rotation-Invariant Descriptor

As mentioned in Section 2.1, building a rotation-invariant local descriptor without assigning a dominant orientation is an effective method to improve the matching performance. Inspired by region partition based on the intensity order [13], we first divide the circular neighborhood of each keypoint into several sub-regions according to the improved intensity order. Ratio orientation histograms of each sub-region are then calculated in a local rotation-invariant coordinate system. Finally, the descriptor is constructed by concatenating the ratio histograms of each sub-region in multiple image

neighborhoods. Consequently, they are inherently rotation invariant, and no dominant orientation is required in the proposed local descriptor. The flow chart of descriptor construction is shown in Figure 2.

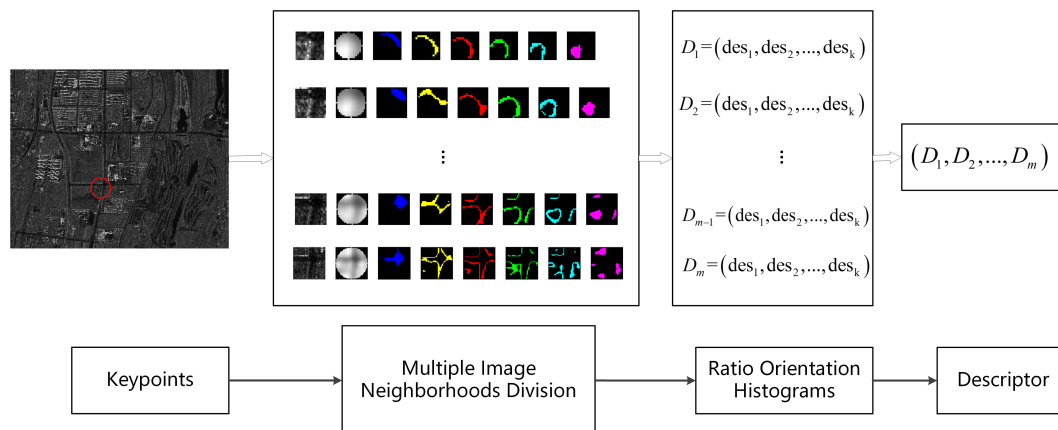


Figure 2. Flow chart of the construction of the proposed descriptor.

2.2.1. Neighborhood Division Based on Improved Intensity Order

SAR images are often acquired at different imaging conditions, such as different times, different polarizations and different viewpoints. Therefore, the same ground scene may appear differently due to differing illumination conditions and sensor sensitivities [20]. Moreover, due to the coherent imaging mode and speckle noise in SAR images, there exist some isolated bright pixels in one image, which can not be found repeatable in another image. Hence, the region division based on the intensity order may fail to correctly partition the neighborhood into similar sub-regions for SAR images. Figure 3 shows the division results of two corresponding neighborhoods in two SAR images. The right-top part of Figure 3 shows the division results based on the intensity order, where different sub-regions are indicated by different colors. We followed the instructions of intensity order pooling in [13]. It can be observed that there exist many unrepeatable fragments in the division results. Hence, it is very difficult to match the two corresponding neighborhoods using the intensity order pooling.

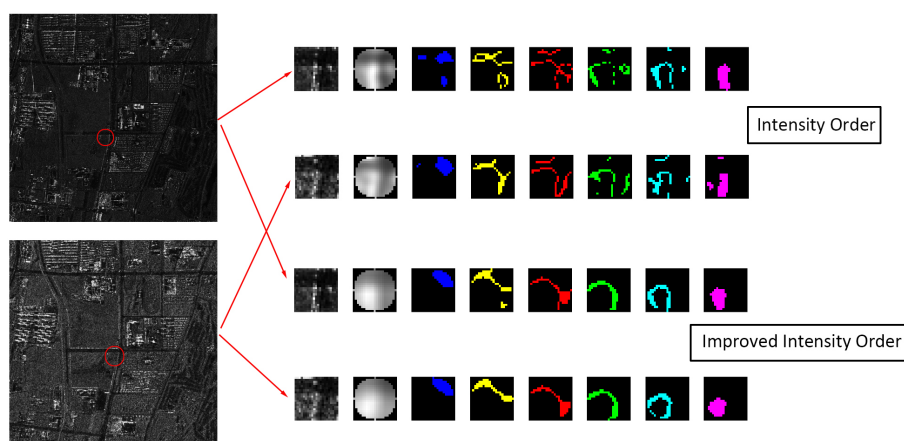


Figure 3. Division results based on the intensity order and the improved intensity order.

Generally, keypoints correspond to locations with significant structural information; their neighborhoods always contain many important feature structures like roads and buildings. However, there also exist small-scale structures caused by high frequency components of speckle noise. Herein, an improved intensity order method is proposed to adapt to SAR images.

First, a median filter is applied to reduce isolated bright pixels. Then, the Rolling Guidance Filter (RGF) proposed by Zhang et al. [26] is utilized to remove small-scale structures while preserving important feature structures. Instead of eliminating speckle noise in the raw image, we aim at searching the large-scale feature structures and effectively dividing the neighborhoods. Compared with the Gaussian filter and other filters specifically designed for SAR images, RGF achieves real-time performance and produces better results in separating different scale structures. Finally, the circular neighborhood is divided according to the intensity order after the two filters. We denote that $P = \{p_1, p_2, \dots, p_h\}$ is the circular neighborhood with h points after the two filters. $I(p_i)$ represents the intensity of the point p_i . The points are first non-descending sorted, and the index is $\{d(i), i = 1, 2, \dots, h : I(p_{d(1)}) \leq I(p_{d(2)}) \leq \dots \leq I(p_{d(h)})\}$. Then, h points are divided into k groups as:

$$P_j = \{p_i \in P : I(t_{j-1}) < I(p_i) \leq I(t_j)\}; t_j = p_{d(\lceil hj/k \rceil)}, j = 1, 2, \dots, k, t_1 = 0. \quad (1)$$

where k groups denote k sub-regions.

The partition result based on the improved intensity order is shown in the right-bottom part of Figure 3. It can be observed that the corresponding sub-regions are very similar, resulting in a high possibility to correctly match the two keypoints. Compared with the result based on the intensity order, it is more reliable for SAR images. The sub-regions are divided by the intensity orders of a circular neighborhood, hence, they are invariant to illumination change and rotation.

2.2.2. Rotation-Invariant Ratio Orientation Histogram

In this section, the rotation-invariant ratio orientation histogram is proposed to construct descriptors for each sub-region. Here, we first build a rotation-invariant coordinate system for each keypoint. We assume that X is a keypoint and P is its circular neighborhood, shown in Figure 4a. For one sample point X_i in the neighborhood, a local coordinate system is established by setting the direction $\vec{XX_i}$ as the x -axis and the direction perpendicular to the $\vec{XX_i}$ as the y -axis. An example of the rotated coordinate system is shown in Figure 4b, it can be observed that the pixels in the original neighborhood of the sample point X_i remain the same in the rotated neighborhood of the corresponding sample point X_j . Hence, the local coordinate system is rotation invariant, descriptors calculated in this coordinate system are also rotation invariant.

Then, we adopt the Gaussian-Gamma-Shaped (GGS) operator [14] to calculate the ratio orientation histogram. For one sample point, its GGS processing window consists of two parts $W_1(x, y)$ and $W_2(x, y)$, given as follows:

$$\begin{aligned} W_1(x, y) &= \frac{|y|^{\alpha-1}}{\sqrt{2\pi}\sigma_x\Gamma(\alpha)\beta^\alpha} \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{|y|}{\beta}\right)\right), y \geq 0, \\ W_2(x, y) &= \frac{|y|^{\alpha-1}}{\sqrt{2\pi}\sigma_x\Gamma(\alpha)\beta^\alpha} \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{|y|}{\beta}\right)\right), y \leq 0, \end{aligned} \quad (2)$$

where $W_1(x, y)$ and $W_2(x, y)$ are two horizontal windows, (x, y) are the coordinate of the point, σ_x , α and β control the size of the processing window and Γ is the gamma function. Rotating the two windows by an orientation angle θ , the two windows are given as $W_s^\theta(x, y) = W_s(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$, $s = 1, 2$. Figure 4c illustrates two processing windows oriented at $\pi/4$. For a processing window, its local mean is computed by the convolution of the image intensities with the window function. The ratio value of an orientation angle θ is calculated by the ratio of two local means as follows:

$$R_i = \log(\mu_1(x, y, \theta_i) / \mu_2(x, y, \theta_i)); \quad i = 1, 2, \dots, n, \quad (3)$$

where $\mu_s(x, y, \theta_i) = \sum W_s^{\theta_i}(x', y')I(x - x', y - y')$; $s = 1, 2$ and $\theta_i = \frac{\pi}{n}(i - 1)$; $i = 1, 2, \dots, n$, n is the number of directions. Hence, for the sample point X_i , we have obtained eight ratio values. The ratio orientation histograms is then computed by accumulating eight ratio values of points in each sub-region. Since we have divided the circular neighborhood into k parts and the ratio orientation histograms of each part have been computed, the final descriptor of this keypoint is derived as:

$$D = (des_1, des_2, \dots, des_k), \quad des_j = (\sum R_1, \sum R_2, \dots, \sum R_n); j = 1, \dots, k, \quad (4)$$

where k is the number of sub-regions and n is the number of orientations. Actually, some information is lost after the improved intensity order pooling. In order to increase the distinctiveness of the descriptor, multiple image neighborhoods are aggregated. Descriptors of several neighborhoods of different sizes are concatenating to form the final descriptor as follows: (D_1, D_2, \dots, D_m) . Here, m is the number of multiple neighborhoods.

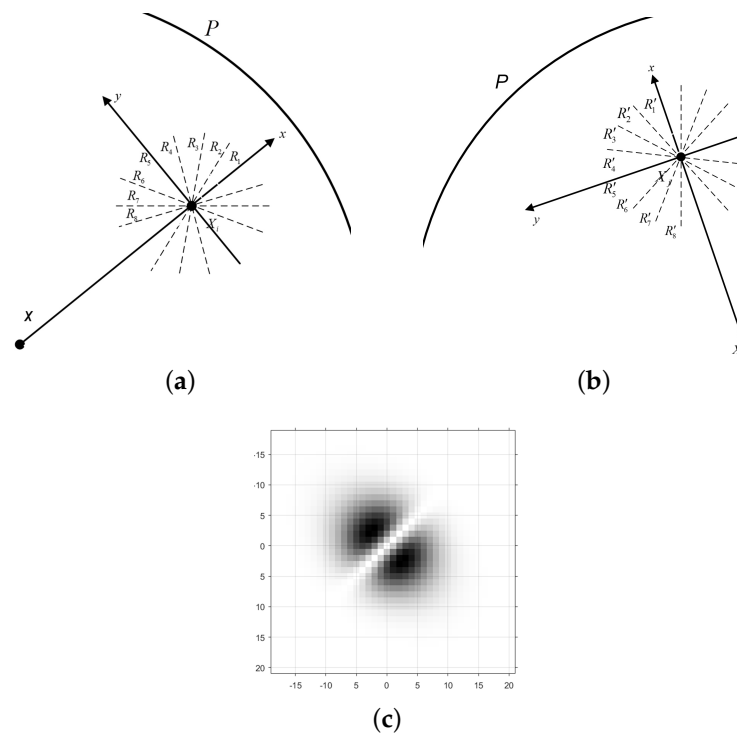


Figure 4. (a) Rotation-invariant local coordinate system. (b) A rotated example of the local coordinate system. (c) A pair of Gaussian-Gamma-Shaped (GGS) windows oriented at $\pi/4$.

3. Experimental Results and Discussion

3.1. Parameter Settings and Datasets

In this section, in order to evaluate the matching performance of the proposed local descriptor, three categories of experiments are operated. The proposed descriptor is compared with three other methods, SIFT, Bilateral Filter Scale-Invariant Feature Transform (BFSIFT) [12], Synthetic Aperture Radar Scale-Invariant Feature Transform (SAR-SIFT) [1]. Since we only focus on the comparisons of descriptors, all of the methods use the same keypoint detection and matching techniques. Herein, keypoints are detected by the SAR-Harris method [1] and matched by the Nearest Neighbor (NN) and Distance Ratio (DR) methods [4]. Parameters of the SIFT, BFSIFT and SAR-SIFT descriptors follow the authors' instructions. For the proposed descriptor, we use multiple neighborhoods with a size of $\{17, 31, 45, 59\}$ to construct the descriptor. Each neighborhood is divided into $k = 6$ parts, and

for each part, $n = 8$ ratio orientation histograms are built. A large orientation number will increase the distinctiveness of the descriptor, while it also results in a heavy computational cost. The GGS processing windows used to calculate the ratio values need to consider the tradeoff of containing enough adjacent pixels and the computational cost. Moreover, the size of the GGS window relates to the scale parameter. Since we do not consider the scale difference in the paper, the size of the GGS window remains the same in all experiments, which is empirically set based on $\{\alpha = 3, \beta = 1.75, \sigma_x = 3.67\}$. The parameters of keypoint detection and matching methods also follow their authors' instructions.

In the first experiment, two satellite images are used; the first is a TerraSAR-X image with a size of 700×700 ; the second is a GaoFen (GF) -3 image with a size of 755×755 . In the second experiment, 16 pairs of TerraSAR-X images are utilized. They are all acquired under the same conditions (StripMap mode, HH polarization, 3-m resolution) with simulated rotations in Beijing City. In the third experiment, three image pairs with complex conditions are used to evaluate the proposed descriptor. These images are obtained under different acquisition conditions, such as polarizations, time and viewpoints, presented in Table 1 and shown in Figure 5. The ground truth transformation model is obtained by manually selecting 20 pairs of control points for each image pair.

Table 1. Image pairs and their characteristics. QPSI, Quad Polarization StripMap; SM, StripMap; FSI, Fine StripMap; DEC, Descending; ASC, Ascending; GF, GaoFen.

Set	Sensor	Mode	Polarization	Date	Direction	Looks	Size
a	Radarsat-2	SM	HH	June 2008	*	1	600×500
a	Radarsat-2	SM	HH	June 2009	*	4	600×500
b	GF-3	QPSI	VV	9 December 2016	DEC	*	552×462
b	GF-3	QPSI	HV	9 December 2016	DEC	*	601×482
c	GF-3	FSI	HH	19 August 2016	DEC	*	564×635
c	GF-3	FSI	HH	20 September 2016	ASC	*	543×629

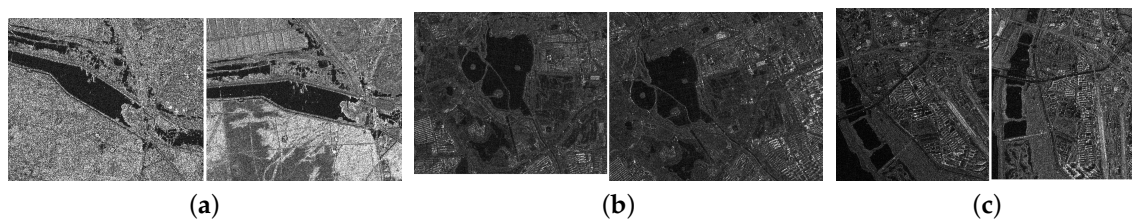


Figure 5. Image pairs. (a) the first; (b) the second; (c) the third.

3.2. Experiments on Rotation Invariance

Herein, we create two experiments to test the rotation invariance of the proposed local descriptor. In the first experiment, we focus on the same images, which only have simulated rotation differences. The raw image is denoted as the reference image; the sensed image is the raw image after rotating by different angles $\beta_k = \frac{k\pi}{8}$, $k = 0, 1, \dots, 15$. We use the bilinear interpolation method in this experiment. Since the calculations of orientation in SIFT and BF-SIFT are the same, we only take SIFT as a representative. We assume that $M(p_1, p_2)$ is a match between a point $p_1(x_1, y_1)$ and a point $p_2(x_2, y_2)$, and the match is a correct match only if $\|H(x_1, y_1) - (x_2, y_2)\|_2 < t$, where H represents the transformation model; $\|\cdot\|_2$ is the Euclidean distance; t stands for a distance threshold. We use the Correct Matches Rate (CMR) to compare the performance of the local descriptors. For a given distance threshold, the CMR is defined as $\text{CMR} = \#cm / \#total$, where $\#cm$ represents the number of correct matches; $\#total$ represents the number of all of the matches. The distance threshold t is used to measure the quality of a correspondence. We set it to five in this experiment.

The curves of CMR versus orientations are shown in Figure 6. A large CMR indicates that more correctly-matched keypoints exist, leading to a more precise transformation model. It can be observed

that the proposed descriptor reaches a large CMR for all orientations. For the SAR-SIFT descriptor, when the angle is in the range of $[100^\circ, 250^\circ]$, its CMRs fall to 0.6. For the SIFT descriptor, its CMRs are close to 0.5 for more than half of the angles. As shown in the experiments in Section 2.1, the calculation of dominant orientation is an error-prone process, and orientation-based methods do not work well for arbitrary positions, the rotation invariance degenerates at some positions. Additionally, since image pairs with larger rotations are more strongly affected by changes due to lighting effects and motion blur [25,27], the expected curve is decreasing as the rotation increases in the range of $[0^\circ, 180^\circ]$. However, we use a bilinear interpolation method to rotate the reference image in this experiment. The interpolating method has an impact on the matching performance, while for 90° , 180° and 270° , the local deformations caused by the interpolating method are slighter than those of other degrees. Consequently, the CMRs of the $90^\circ, 180^\circ, 270^\circ$ for SIFT and SAR-SIFT are higher than those of their adjacent degrees.

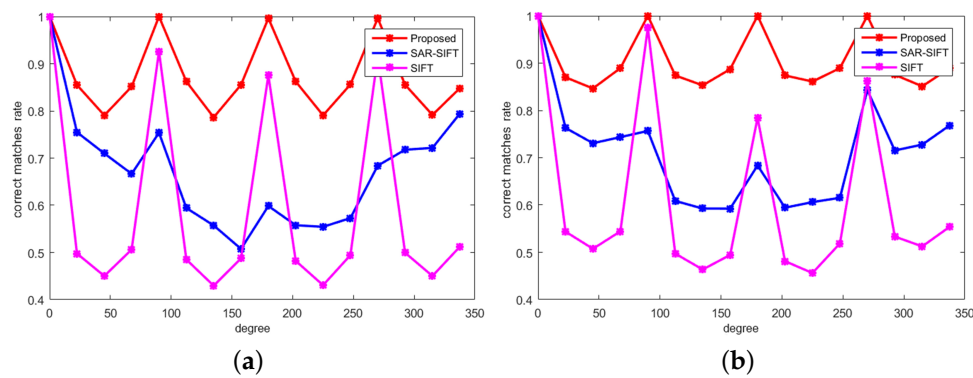


Figure 6. The curves of Correct Matches Rate (CMR) versus angles. (a) the GF-3 image; (b) the TerraSAR-X image.

Hence, for the reference and sensed images with rotation differences, the orientation estimation turns many true corresponding points into misregistrations by their descriptors. Instead of adopting dominant orientation, the proposed descriptor makes use of the improved intensity orders and the rotation-invariant ratio orientation histogram, resulting in more robust rotation invariance.

In the second experiment, Receiver Operator Characteristic (ROC) curves are used to compare the matching performance of the four descriptors. The curve shows the percentage of correctly-matched keypoints against the false alarm rate. The two criteria are given as:

$$Per_{CM} = \frac{\#CM}{\#M_{all}}, \quad false\ alarm\ rate = \frac{\#CM}{\#CM + \#FM}, \quad (5)$$

where $\#CM$ represents the number of correct matches, $\#FM$ represents the number of false matches and $\#M_{all}$ stands for the number of total matches. For different ratio threshold th used in the Nearest Neighbor (NN) and Distance Ratio (DR) method, global ROC curves can be obtained by the Per_{CM} against the *false alarm rate*, shown in Figure 7.

It can be observed that the proposed descriptor gives the best matching performance, followed by the SAR-SIFT algorithm. The performances of BFSIFT and SIFT algorithms are similar. Considering the speckle noise, SAR-SIFT takes advantages of the gradient by ratio and gives better performance than SIFT and BFSIFT on SAR images. However, the aforementioned three SIFT-like algorithms all adopt the dominant orientation assignment. For our proposed descriptor, we adopt the GGS detector to construct the ratio orientation histogram; it is also robust to speckle noise. Moreover, the proposed descriptor combines the information of multiple image neighborhoods to increase the distinctiveness, and it is rotation invariant without relying on a dominant orientation, further improving its robustness.

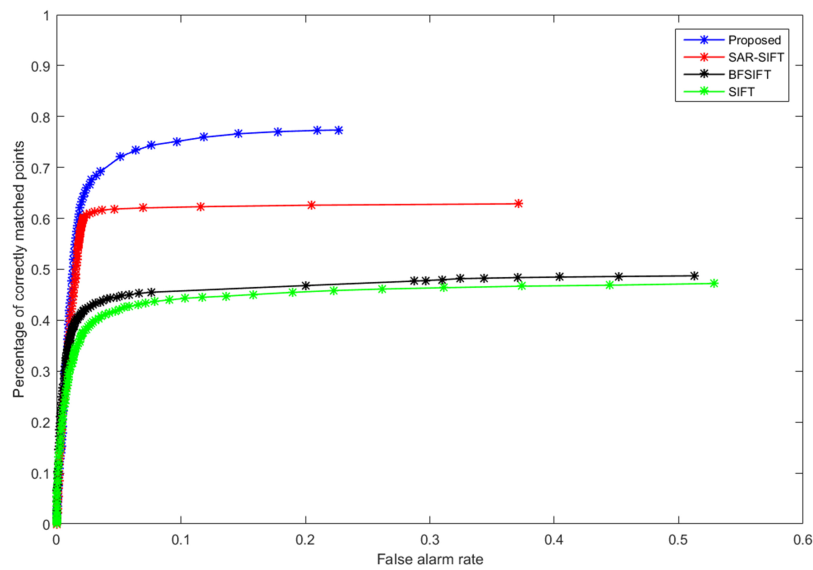


Figure 7. Global ROC curves of the proposed descriptor, SAR-SIFT, BFSIFT and SIFT.

3.3. Experiments on Satellite SAR Images with Complex Conditions

In these experiments, three image pairs obtained under complex situations are utilized to further evaluate the performance of the four descriptors. We adopt the RANSAC method to remove the false matches for all of the comparative methods. For two SAR images that have the same resolution, we skip the scale space construction in order to increase the matching process. The matching performance is quantitatively evaluated by the Root Mean Square Error (RMSE) and CMR. The RMSE can be computed as $\zeta_{\text{rmse}} = \frac{1}{N_{\text{corr}}} \sum_{i=1}^{N_{\text{corr}}} \|H(x_1^i, y_1^i) - (x_2^i, y_2^i)\|_2$, where (x_1^i, y_1^i) and (x_2^i, y_2^i) are the coordinates of the i th matched pair and N_{corr} is the number of correctly-matched keypoints, and the Standard Deviation (SD) and Maximum Error (ME) are also presented. Small RMSE denotes that the accuracy of matching performance is high. Comparisons of the four descriptors are presented in Table 2. Matching results of the second image pair are shown in Figure 8, where yellow lines denote correctly-matched correspondences, and red lines denote misregistrations.

Table 2. Comparisons of SAR-SIFT, BFSIFT, SIFT and the proposed method.

	Set	SAR-SIFT	BFSIFT	SIFT	Proposed
CMR	a	15/44	29/73	24/76	30/40
RMSE (pixel)	a	1.22	1.28	1.63	0.89
SD	a	0.6545	1.0818	1.4436	0.4816
ME	a	2.5840	2.3780	4.4502	1.9676
CMR	b	17/44	7/45	9/57	21/41
RMSE (pixel)	b	1.91	2.57	2.78	1.29
SD	b	1.1503	1.5380	1.5040	0.7455
ME	b	2.8800	4.2722	5.6833	2.7870
CMR	c	13/56	9/66	7/55	9/54
RMSE (pixel)	c	1.45	2.89	2.34	1.78
SD	c	0.6510	1.7002	1.2165	1.1036
ME	c	2.0994	6.5029	4.6881	3.9308

It can be observed from Table 2 that the proposed descriptor gives a better matching performance than the three other descriptors on the first and second image pairs, followed by SAR-SIFT. As shown in Figure 8, the proposed descriptor yields the highest number of correctly-matched keypoints,

resulting in a more precise transformation model. Compared with SIFT, BFSIFT only replaces the Gaussian filter with the bilateral filter; it gives better matching performance in regions with edge structures, whereas SIFT yields better results in regions with blob structures. Since there are many edge structures in the first image pair, BFSIFT gives better results than SIFT. However, the third image pair describes a dense urban area. Due to the side-looking mechanism of SAR sensors, urban areas are corrupted with geometric distortions, such as layover and foreshortening [28]. Strongly affected by distortions, neighborhood division in the proposed descriptor becomes unstable, resulting in an increasing number of misregistrations, presented in Table 2. How to make the proposed descriptor more adaptable to local distortions will be studied in the future.

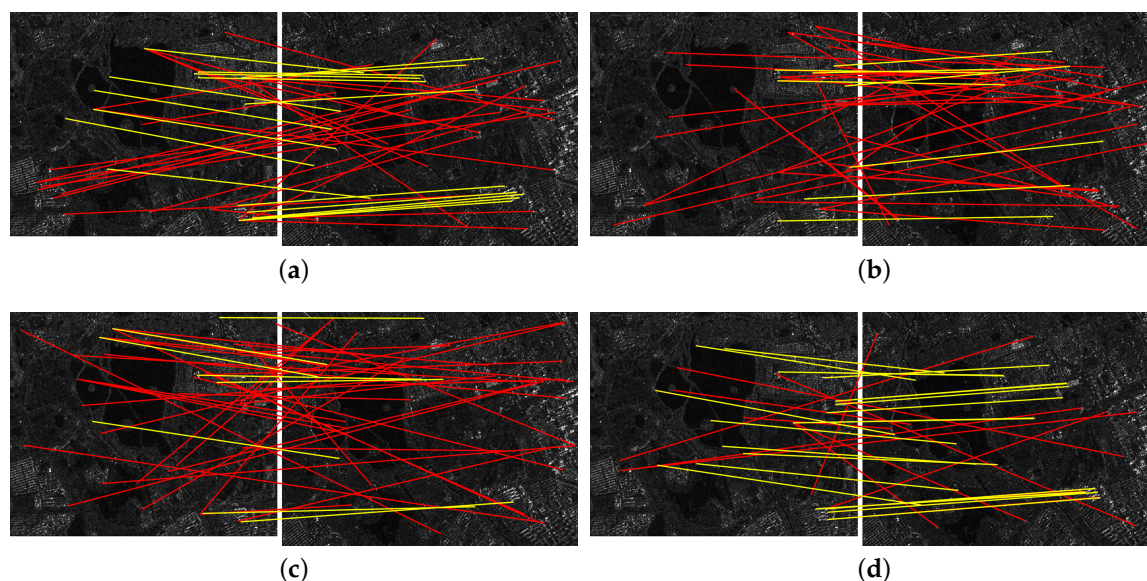


Figure 8. Matching results on second image pair. (a) SAR-SIFT; (b) BFSIFT; (c) SIFT; (d) proposed.

4. Conclusions

In this paper, an advanced rotation-invariant descriptor for SAR image registration is proposed. Aiming at achieving rotation invariance without assigning a dominant orientation, we first divide the circular neighborhood into several sub-regions based on the improved intensity orders. Then, a rotation-invariant local coordinate system is built for each keypoint. The ratio orientation histogram of each sub-region is calculated by the GGS ratio values of different directions in this coordinate system. Moreover, multiple image neighborhoods are aggregated to increase the distinctiveness of the descriptor. The proposed descriptor is composed of the concatenation of the ratio orientation histograms of each sub-region in multiple neighborhoods. Experimental results show that the proposed descriptor has achieved a robust rotation invariance and yields a superior matching performance for SAR image registration.

Acknowledgments: The authors would like to thank Zelian Wen from Xidian University for providing the source code of SAR-SIFT algorithm and comments from the anonymous reviewers.

Author Contributions: Yuming Xiang conceived of and performed the experiments. Feng Wang and Hongjian You supervised the research. Ling Wan carried out the comparative analysis. Yuming Xiang and Feng Wang drafted the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dellinger, F.; Delon, J.; Gousseau, Y.; Michel, J.; Tupin, F. SAR-SIFT: A SIFT-like algorithm for SAR images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 453–466.
2. Sui, H.; Xu, C.; Liu, J.; Hua, F. Automatic optical-to-SAR image registration by iterative line extraction and Voronoi integrated spectral point matching. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6058–6072.
3. Wang, Z.; Zhang, J.; Zhang, Y.; Zou, B. Automatic registration of sar and optical image based on multi-features and multi-constraints. In Proceedings of the 2010 IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS), Honolulu, HI, USA, 25–30 July 2010; pp. 1019–1022.
4. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
5. Huang, L.; Li, Z. Feature-based image registration using the shape context. *Int. J. Remote Sens.* **2010**, *31*, 2169–2177.
6. Zhao, J.; Gao, S.; Sui, H.; Li, Y.; Li, L. Automatic registration of SAR and optical image based on line and graph spectral theory. In Proceedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Suzhou, China, 14–16 May 2014; p. 377.
7. Ma, J.; Zhou, H.; Zhao, J.; Gao, Y.; Jiang, J.; Tian, J. Robust feature matching for remote sensing image registration via locally linear transforming. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6469–6481.
8. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395.
9. Rabin, J.; Delon, J.; Gousseau, Y.; Moisan, L. MAC-RANSAC: A robust algorithm for the recognition of multiple objects. In Proceedings of the Fifth International Symposium on 3D Data Processing, Visualization and Transmission (3DPTV 2010), Paris, France, 17–20 May 2010; p. 051.
10. Schwind, P.; Suri, S.; Reinartz, P.; Siebert, A. Applicability of the SIFT operator to geometric SAR image registration. *Int. J. Remote Sens.* **2010**, *31*, 1959–1980.
11. Fan, B.; Huo, C.; Pan, C.; Kong, Q. Registration of optical and SAR satellite images by exploring the spatial relationship of the improved SIFT. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 657–661.
12. Wang, S.; You, H.; Fu, K. BFSIFT: A novel method to find feature matches for SAR image registration. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 649–653.
13. Fan, B.; Wu, F.; Hu, Z. Aggregating gradient distributions into intensity orders: A novel local image descriptor. In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011; pp. 2377–2384.
14. Shui, P.L.; Cheng, D. Edge detector of SAR images using Gaussian-Gamma-shaped bi-windows. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 846–850.
15. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up robust features. In Proceedings of the European Conference on Computer Vision, Graz, Austria, 7–13 May 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417.
16. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.
17. Xia, G.S.; Delon, J.; Gousseau, Y. Shape-based invariant texture indexing. *Int. J. Comput. Vis.* **2010**, *88*, 382–403.
18. Dai, X.; Khorram, S. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2351–2362.
19. Yasein, M.S.; Agathoklis, P. Automatic and robust image registration using feature points extraction and Zernike moments invariants. In Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology, Athens, Greece, 18–21 December 2005; pp. 566–571.
20. Wong, A.; Clausi, D.A. ARRSI: Automatic registration of remote-sensing images. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1483–1493.
21. Yang, W.; Dai, D.; Triggs, B.; Xia, G.S. SAR-based terrain classification using weakly supervised hierarchical Markov aspect models. *IEEE Trans. Image Process.* **2012**, *21*, 4232–4243.
22. Hu, F.; Xia, G.S.; Hu, J.; Zhang, L. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens.* **2015**, *7*, 14680–14707.

23. Yang, W.; Yin, X.; Xia, G.S. Learning high-level features for satellite image classification with limited labeled samples. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4472–4482.
24. Hu, F.; Xia, G.S.; Wang, Z.; Huang, X.; Zhang, L.; Sun, H. Unsupervised feature learning via spectral clustering of multidimensional patches for remotely sensed scene classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2015–2030.
25. Liu, K.; Skibbe, H.; Schmidt, T.; Blein, T.; Palme, K.; Brox, T.; Ronneberger, O. Rotation-invariant HOG descriptors using fourier analysis in polar and spherical coordinates. *Int. J. Comput. Vis.* **2014**, *106*, 342–364.
26. Zhang, Q.; Shen, X.; Xu, L.; Jia, J. Rolling guidance filter. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 815–830.
27. Gauglitz, S.; Turk, M.; Höllerer, T. Improving Keypoint Orientation Assignment. In Proceedings of the British Machine Vision Conference (BMVC), Dundee, UK, 29 August–2 September 2011; pp. 1–11.
28. Salentinig, A.; Gamba, P. Combining SAR-based and multispectral-based extractions to map urban areas at multiple spatial resolutions. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 100–112.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).