EEKHI–SAR–SIFT: Edge Enhancement and Keypoint Homogeneous Improved SAR–SIFT Framework Based on Unbiased Difference Ratio Edge Detector Strategy

Zhonghua Hong, *Member, IEEE*, Yu Lu, Yongsheng Geng, Xiaohua Tong, *Senior Member, IEEE*, Shijie Liu, Ruyan Zhou, Haiyan Pan, Yun Zhang, Yanling Han, Jing Wang, and Shuhu Yang

***Abstract*—Keypoint matching plays a vital role in the realm of synthetic aperture radar (SAR) image processing, serving as a crucial component within this domain. Another key fact to remember keypoint matching is a crucial step in change detection and image stitching. In the context of SAR images, a scale-invariant feature transformation (SIFT)-based approach, known as SAR–SIFT, presents a notable advantage by mitigating the impact of speckle noise; however, it cannot yield accurate edge information, and the resulting keypoints are nonuniformly distributed. We propose an edge enhancement and homogeneous spatial key point improved SAR–SIFT framework based on an unbiased difference-ratio (UDR) edge detector (called EEKHI–SAR–SIFT) to solve the above problems. The algorithm relies on the characteristics of edge unbiased localization and constant false alarm rate of UDR edge extraction to reduce the extracted wrong corner information as well as enhance the extraction exactness of keypoints. In addition, adaptive nonmaximum suppression (ANMS) method is applied to homogenize the dense keypoints with a large initial number that are gained by means of the EEKHI–SAR–SIFT algorithm and reduce their local clustering. Finally, a descriptor construction strategy that retains multiscale information is adopted to improve the descriptor uniqueness. Tests using multiple sets of SAR image data from different satellites (Gaofen-3, RADARSAT, and Sentinel-1A) demonstrate that the efficacy of the proposed EEKHI–SAR–SIFT algorithm reduces the root mean square error is about 1–2 pixel lower than the final result of the original SAR–SIFT algorithm.**

***Index Terms*—Nonmaximal suppression point detection, ratio of exponentially weighted average (ROEWA) detector, SAR–SIFT, synthetic aperture radar imagery, unbiased difference-ratio (UDR) detector.**

Manuscript received 1 April 2024; revised 9 June 2024 and 12 July 2024; accepted 1 August 2024. Date of publication 6 August 2024; date of current version 15 August 2024. This work was supported in part by the National Key R&D Program of China under Grant 2018YFB0505400, and in part by the National Natural Science Foundation of China under Grant 41871325. *(Corresponding author: Xiaohua Tong.)*

Zhonghua Hong, Yu Lu, Yongsheng Geng, Ruyan Zhou, Haiyan Pan, Yun Zhang, Yanling Han, Jing Wang, and Shuhu Yang are with the College of Information Technology, Shanghai Ocean University, Shanghai 201306, China, and also with the Key Laboratory of Fisheries Information, Ministry of Agriculture, Shanghai Ocean University, Shanghai 201306, China (e-mail: zhhong@shou.edu.cn; m210901466@shou.edu.cn; m180701065@shou.edu.cn; ryzhou@shou.edu.cn; hy-pan@shou.edu.cn; y-zhang@shou.edu.cn; ylhan@shou.edu.cn; wangjing@shou.edu.cn; shyang@shou.edu.cn).

Xiaohua Tong and Shijie Liu are with the College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China (e-mail: xhtong@tongji.edu.cn; liusjtj@tongji.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3438796

I. Introduction

IMAGE registration of synthetic aperture radar (SAR) images is extensively used in change detection [1], image stitching [2], image retrieval [3] and interferometric SAR technology, and SAR image keypoint extraction is an important basis for these applications. Therefore, to ensure the accuracy of the above image processing results, it is crucial to design a high-precision image matching process [4]. Currently, it is hard to extract unique, robust, and uniformly dispensed features from SAR images because of their remarkable intensity variations and geometric differences and the unique multiplicative speckle noise problem [5].

Depending on the matching method, image registration can be divided into two categories, feature-based matching (FBM) [6] and area-based matching (ABM) [7]. An ABM method directly matches the keypoints in a search area based on the grayscale feature of the image [8]. The most representative ABM method is the normalized cross-correlation (NCC) method [9], which matches by calculating the correlation of the image window to be matched. The NCC method is commonly used because of its robustness to linear radiation transformation while effectively diminishing the influence of illumination on the matching accuracy. Although NCC is extensively used in image matching, it cannot deal with nonlinear image relationships [10]. Other region-based matching methods include the sequential similarity detection algorithm [11] and the mutual information [12].

The FBM algorithm obtains the correspondence between images by extracting and describing the keypoints with a certain texture. The most classic FBM algorithm is scale-invariant feature transformation (SIFT) [13]. The SIFT algorithm extracts key point features based on the local gradient direction histogram [14], which is widely used in image registration due to its characteristics of rotation, illumination, and scale invariance [15].

In the past research, it was found that the SIFT matching algorithm has problems such as large image size, aggregation of matching points and large mismatch ratio [16]. SAR-SIFT used the ratio of exponentially weighted averag (ROEWA) [17] filtering for each scale-space in the constructed scale-space to ensure that the obtained scale-space images have CFAR properties. In terms of descriptor construction, a gradient location

and orientation histogram (GLOH) [18] is used to improve the robustness and uniqueness of feature descriptors.

The SAR–SIFT algorithm leads to corresponding improvements in the above three aspects and successfully applies the SIFT algorithm to SAR image data. Bangsong et al. [19] proposed the USAR-SIFT method, using the triangulation model of the traditional data model to optimize the spatial distribution of the keypoints in SAR images. On the contrary, this algorithm also has the problems of insensitivity to image edge information. Paul and Patti [20] proposed the I–SAR–SIFT algorithm, in the process of SAR–SIFT feature description, the multiscale feature descriptor construction method is adopted, which not only improves the uniqueness of feature descriptors, but also improves the accuracy of matching points. However, this method has a drawback, the matching points are mainly concentrated in the clearly textured areas [21]. Fan et al. [22] proposed the use of sparse representation technology and designed a new feature descriptor for matching, which can significantly improve matching performance and accuracy. According to the inherent characteristics of optical and SAR images, Xiang et al. [23] improved the key point detection, direction allocation, and other aspects of scale space to improve the registration accuracy. Hong et al. [24] proposed the ROEWA–Blocks algorithm by introducing the ROEWA operator in the SAR–SIFT algorithm. The ROEWA–Blocks algorithm has been used to solve the problem of abrupt dark blocks in SAR images, and concentric circles have been used to save multiscale image features. Xiang et al. [25] estimate scaling and rotation differences of SAR imagery from metadata and use a novel matching framework to speed up the flow of matching. However, none of these concepts can deal with the barrier of inaccurate edge localization in an SAR image feature extraction algorithm. Xiang et al. [26] proposed a stable keypoint detector based on feature intersection, and designed cross-stage partial Siamese network to enhance the uniqueness of descriptors. Finally, an instance mining algorithm was employed to improve matching efficiency, resulting in satisfactory matching outcomes. However, the method is unable to capture high-resolution texture details, leading to some deviation in the localization of the obtained matching points. Xiang et al. [27] employed a joint filter of refined-lee filtering and polarization whitening filtering, which, while limiting the impact of speckle noise, preserves the details of high-resolution imagery. A novel Siamese multiscale attention network was designed for constructing multiscale descriptors, ultimately achieving precise SAR image registration. However, the distribution of the matching points obtained by this method is not uniform, with some local areas lacking matching points in the experimental results. In contrast, the algorithm presented in this article achieves a matching point accuracy in Gaofen-3 satellite imagery that is comparable to the aforementioned algorithms. However, the adaptive nonmaximum suppression (NMS) method employed in this study yields a uniformly distributed set of matching points, and the adaptive NMS method chosen in this article can yield a uniformly distributed set of matching points. In the aforementioned methods, only the acquisition of feature points unaffected by image noise and illumination variations can be ensured. However, the edge texture information in the filtered image becomes blurred, leading to some feature points obtained by the SAR–SIFT algorithm deviating from the texture edges of the image. These feature points do not fall within the scope of true feature points. We have chosen to improve with the unbiased difference-ratio (UDR) edge detector, which maintains high-resolution texture details while significantly reducing the impact of speckle noise. Compared to other algorithms, the feature points obtained in this article are closer to the image edges and can effectively reduce the number of such feature points.

This approach is innovative in the following respects.

1) This article employs UDR filtering in the scale space, replacing the ROEWA filtering used by SAR–SIFT. The UDR filtering, which combines differential and ratio operations, not only mitigates the speckle noise in SAR imagery but also more accurately preserves image details, thereby capturing superior textural and structural features of the images.
2) The feature points obtained in the SAR–SIFT algorithm exhibit spatial clustering and redundancy. Traditional NMS algorithms can reduce local aggregation but do not ensure the acquisition of globally distributed feature points. Therefore, this article opts for a fast adaptive NMS algorithm, which adjusts dynamic thresholds to ensure that the selected feature points are distributed across the entire image. Furthermore, to enhance algorithmic efficiency, CPU multithreading is utilized for accelerated processing.
3) Building upon the descriptor construction of SAR-SIFT, this article adopts a 45° bin partitioning strategy and combines it with multiscale concentric circular support regions to construct descriptors, significantly enhancing the uniqueness of the descriptors.

The rest of this article is set as follows. In Section II we introduce some improvements to the algorithm. In Section III, we present qualitative and quantitative comparison results between our algorithm and other algorithms. Ultimately, in Section IV, a summary of the algorithm in this article is made.

II. Improved Proposals

The high-precision and uniform keypoint extraction method exhibited in Fig. 1. The specific explanation is as follows: the stage A replaces the ROEWA operator in the SAR–SIFT algorithm with UDR detector to obtain high-quality SAR image gradient maps. The stage B uses an adaptive NMS method to optimize the keypoints in different regions. The stage C saves multiscale information in concentric circles and subsequently constructs feature descriptors.

A. UDR Edge Detector for Constructing Multiscale Space

The ROEWA operator used in the SAR–SIFT algorithm only considers ensuring the CFAR of the image, without taking into account the precise localization of edge information within the image. This article introduces an UDR detector that combines ratio and difference operations. This detector possesses the characteristic of CFAR in the ratio operation and precise edge localization in the difference operation. In this article, the

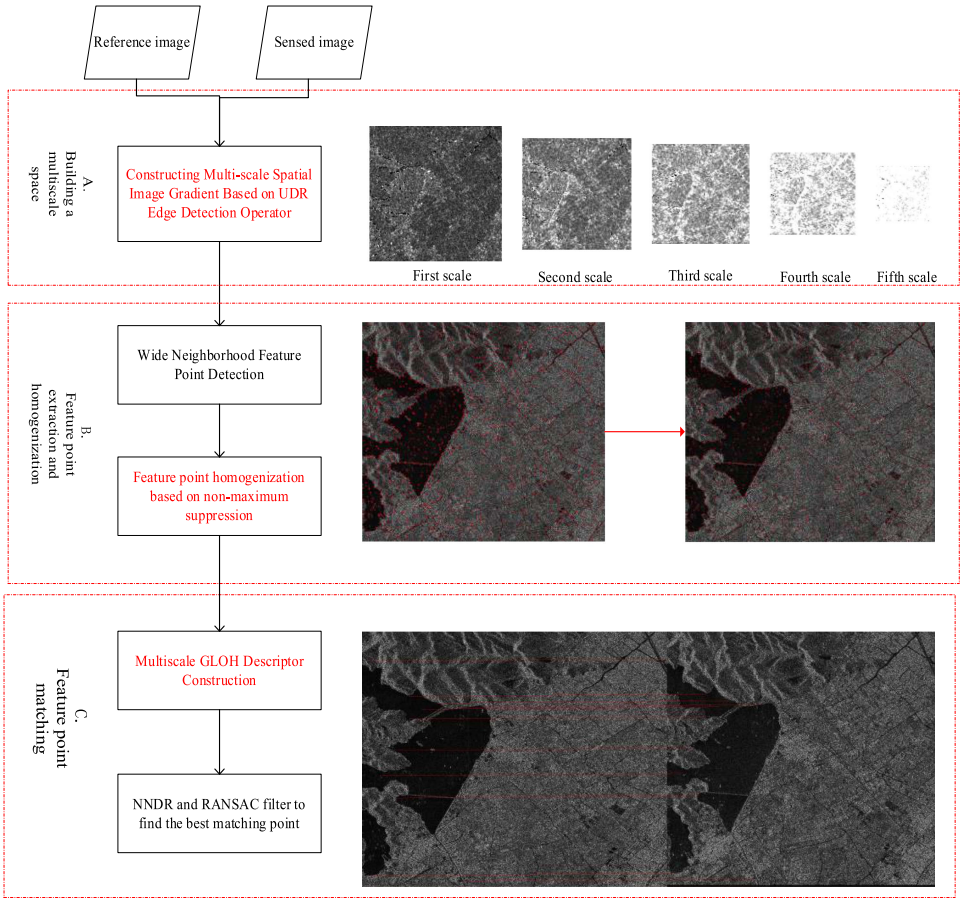


Fig. 1. Algorithm flowchart of this article (NNDR: nearest neighbor distance ratio and RANSAC: random sample consensus).

ROEWA operator used in the SAR–SIFT algorithm is replaced with the aforementioned UDR detector in the scale space. This replacement effectively reduces the number of false keypoints in the SAR–SIFT algorithm, thereby achieving the extraction of high-precision keypoints.

The EEKHI–SAR–SIFT algorithm proposed in this article improves the UDR detector to adapt to the SAR–SIFT scale space and uses detection filters, which are represented by $h_{\text{name}}(x, y)$. Superscripts "$\|$" and "$\perp$" represent parallel and vertical detection filters, respectively. The one-dimensional (1-D) vertical filter can be expressed as follows:

$$\begin{aligned} h_{\text{UDR}}^{\perp}(x) &= |x|^{\alpha-1}\exp\left(-\frac{|x|}{\beta}\right)\varepsilon(x) \\ &= h_{\text{UDR}}^{\perp}(x) + h_{\text{UDR}}^{\perp}(-x) \\ h_{\text{UDR}}^{\perp}(x) &= |x|^{\alpha-1}\exp\left(-\frac{|x|}{\beta}\right)\varepsilon(x)/\rho\eta^{i} \\ h_{\text{UDR}}^{\perp}(-x) &= |x|^{\alpha-1}\exp\left(-\frac{|x|}{\beta}\right)\varepsilon(-x)/\rho\eta^{i}. \end{aligned} \tag{1}$$

In (1), parameters $\alpha > 1$ and $\beta > 0$ control the peak value and vertical filter, $\varepsilon(\cdot)$ respectively represents unit step function, ρ the initial layer scale in the SAR–SIFT feature extraction (default is 2), η the adjacent two-layer scale ratio (default $\sqrt[3]{2}$), and i the scale space layer number of the SAR–SIFT algorithm (default is 5). The 1-D parallel filter can be expressed as follows:

$$h_{\text{UDR}}(y) = \begin{cases} 1/\rho\eta^{i} & |y| \leqslant l_{\|} \\ \exp\left(\dfrac{-(|y|-l_{\|})^{2}}{2\sigma_{\|}^{2}}\right)/\rho\eta^{i} & |y| \geqslant l_{\|} \end{cases} \tag{2}$$

where parameters $l_{\|}$ and $\sigma_{\|}$ represent the flatness and decay degree of the bell-shaped function, respectively. A 2-D edge detector with UDR detector is written as follows:

$$\begin{aligned} h_{\text{UDR}}(x) &= h_{\text{UDR}}^{\perp}(|x|)\, h_{\|}(y) \\ &= h_{\text{UDR}}(+x, y) + h_{\text{UDR}}(-x, y). \end{aligned} \tag{3}$$

Parameter $+x$ represents the left sides of the UDR detector and parameter $-x$ represents the right sides. In this study, this UDR detector replaces the original ROEWA operator in the SAR–SIFT algorithm. The objectives are to adapt to the multi-scale space requirements of SIFT, decrease the influence of speckle noise in SAR images, and obtain unbiased image edge location information.

To match an edge with the direction of the UDR detector, the latter needs to be rotated. The UDR detector direction is expressed as follows:

$$h_{\text{UDR}}^{\theta}(x) = h_{\text{UDR}}^{\theta}(+x, y) + h_{\text{UDR}}^{\theta}(-x, y)$$

$$h_{\mathrm{UDR}}(+x,y)=h\left((+x)\cos\theta-y\sin\theta,(+x)\sin\theta+y\cos\theta\right)$$
$$h_{\mathrm{UDR}}(-x,y)=h\left((-x)\cos\theta-y\sin\theta,(-x)\sin\theta+y\cos\theta\right). \tag{4}$$

In (4), parameter θ is the angle of the desired rotation.

Many different kinds of feature detectors have been proposed in optical images, such as Harris–Laplace [29] and Harris-affine [30]. When these feature detectors in optical images are directly used for SAR image matching, the effect is not satisfactory. Therefore, ROEWA filtering is used to process the image in SAR–SIFT to reduce the influence of speckle noise, and the gradient by ratio (GR) method is used to calculate the gradient magnitude and gradient direction of an image. This method is robust to speckle noise, particularly for SAR imagery. Specifically, the logarithm of the area-weighting ratios of the left and right areas in the horizontal direction and the upper and lower areas in the vertical direction of the ROEWA operator are employed to try to solve the gradient magnitude and direction of a pixel point. Moreover, (6) is used to calculate the horizontal and vertical gradients. Therefore, according to the principle of GR calculation, horizontal and vertical rotation angles are input in the UDR detector. In addition, the horizontal and vertical gradients are calculated using the GR algorithm formula as follows:

$$U_{\mathrm{horizontal}}=\frac{h^{0^\circ}_{\mathrm{UDR}}(+x,y)}{h^{0^\circ}_{\mathrm{UDR}}(-x,y)},\ U_{\mathrm{vertical}}=\frac{h^{90^\circ}_{\mathrm{UDR}}(+x,y)}{h^{90^\circ}_{\mathrm{UDR}}(-x,y)}$$
$$Gx,\alpha=\log\left(U_{\mathrm{horizontal}},\sigma\right)$$
$$Gy,\alpha=\log\left(U_{\mathrm{vertical}},\sigma\right). \tag{5}$$

Moreover, the gradient magnitude and direction are calculated using the following:

$$Gm,\sigma=\sqrt{(Gx,\sigma)^2+(Gy,\sigma)^2}$$
$$Go,\sigma=\arctan\left(\frac{Gx,\sigma}{Gy,\sigma}\right) \tag{6}$$

where Gm,σ is used to describe the gradient magnitude and Go,σ is used to describe the gradient direction. And α in (5) is the weight parameter used to calculate the local mean, whereas σ in (6) refers to the scale space where the UDR detector is located. Because the SARHarris method cannot effectively determine the feature scale parameters, it cannot yield the feature positions in the scale space

Because the SAR–Harris method cannot effectively determine the feature scale parameters, it cannot yield the feature positions in the scale space. Yu et al. [2] combined the Harris–Laplace and ROEWA methods to obtain accurate positions and corresponding scale information of keypoints. Therefore, this article proposes to use the Harris–Laplace algorithm to achieve features in SAR images. The next two formulas are the Harris–Laplace matrix and a corner detector based on the UDR detector are formulated as follows:

$$M(x,y,\sigma_n)$$

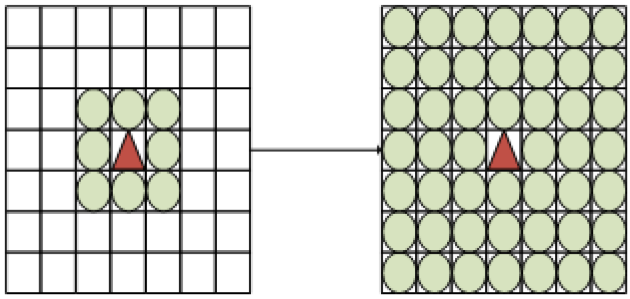


Fig. 2. Spatial domain expansion map of candidate keypoints.

$$=\sigma_n^2\cdot g\left(\sqrt{2}\sigma_n\right)*\begin{bmatrix}(Gx,\alpha)^2 & (Gx,\alpha)(Gy,\alpha)\\ (Gx,\alpha)(Gy,\alpha) & (Gy,\alpha)^2\end{bmatrix}$$
$$R(x,y,\sigma_n)$$
$$=\det\left(M(x,y,\sigma_n)\right)-d\cdot\mathrm{tr}(M(x,y,\sigma_n))^2 \tag{7}$$

where σ_n denotes the nth scale space, $g(\sqrt{2}\sigma_n)$ represents a 2-D Gaussian function with a standard deviation of $\sqrt{2}\sigma_n$, and d is an arbitrary parameter (default is 0.04).

B. Keypoint Detection and Homogenization

Because the SAR–Harris method cannot effectively determine the feature scale parameters, it cannot yield the feature positions in the scale space.

In the SAR–SIFT algorithm, detection of a candidate keypoint only compares the 3×3 spatial neighborhood of the current layer to determine if it is an extreme point. When applying the SAR–SIFT algorithm to extract keypoints, the majority of keypoints cluster in areas with large pixel grey values, and the keypoints are spaced within two pixels. These keypoints with relatively close spatial positions may lead to more mismatched points with the subsequent nearest neighbor algorithm. This study solves the above problems by expanding the comparison range of candidate keypoints. As shown in Fig. 2, the comparison is performed between 5×5 and 7×7 spatial neighborhoods in the same scale space. If a candidate keypoint is an extreme point, it is retained; otherwise, it is removed. By expanding the range of the domain comparison, extraction of the keypoints that are extremely close in space can be limited and the phenomenon of keypoint clustering in grey areas can be reduced. Moreover, the recognizability of corner points can be increased, and improved keypoint information can be provided for the subsequent keypoint matching.

After the above operations, although keypoints with a more uniform spatial distribution can be obtained compared with the SAR–SIFT algorithm, numerous calculated keypoints are still clustered and redundant. Therefore, it is necessary to eliminate a portion of the redundant and locally clustered initial feature points. The elimination process typically employs the NMS algorithm. Traditional NMS algorithms retain only the feature points with the highest response within each neighboring region, discarding all others. The size of the neighboring region is

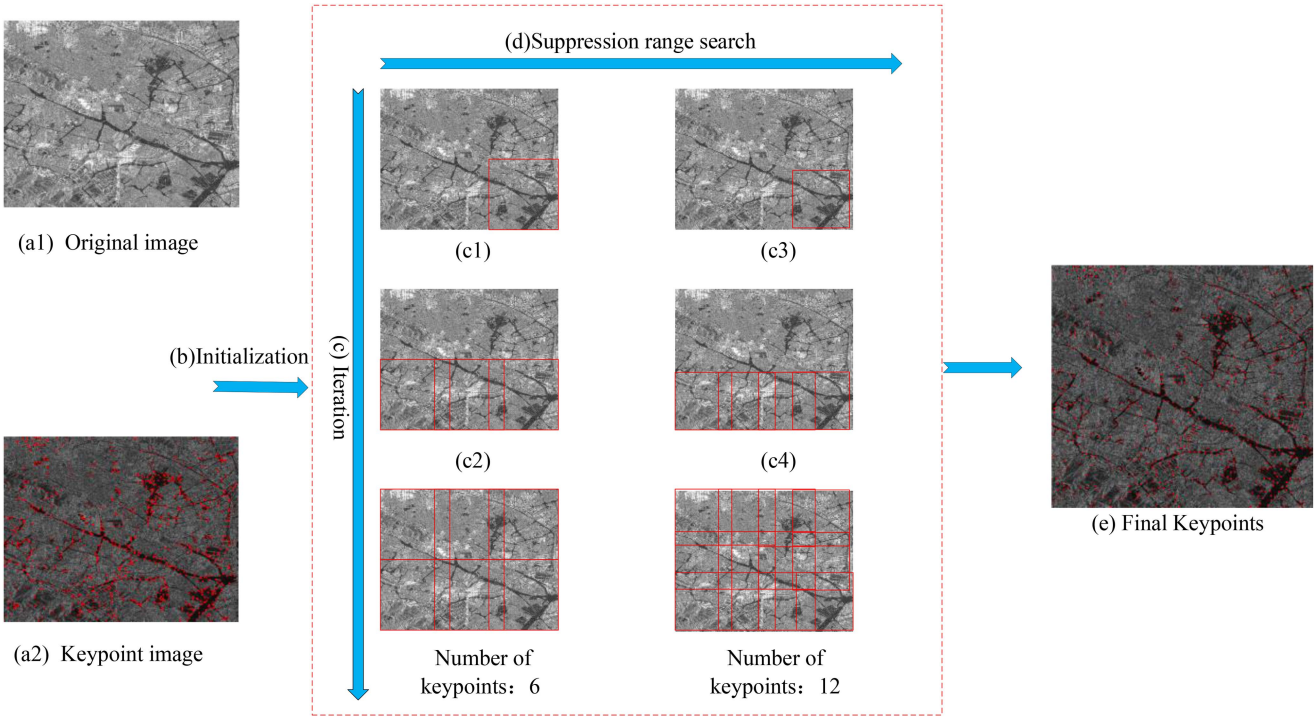


Fig. 3. SSC algorithm flow. (a1) is original image. (a2) is keypoint image. (b) is initialization of entire process. (c) represents obtaining of strong response points throughout image with square suppression. (d) represents binary search to determine square side length d. (e) represents final uniform keypoints.

usually manually set. This article opts to improve the aforementioned algorithm by employing a square suppression technique.

Fig. 3 shows the general flow of the ANMS algorithm, compared to the traditional NMS algorithm, the fast adaptive NMS algorithm employed in this article introduces a dynamic threshold mechanism instead of relying on a fixed threshold. This dynamic thresholding allows for a more uniform distribution of the selected feature points across the image, enhancing the detection performance by ensuring a balanced representation of features throughout the scene. a1 is the premier image and a2 is the initial keypoints after keypoint extraction from the initial image. In the b process, sorting is performed according to the intensities of the keypoints, and the side length of the square area is initialized, NMS algorithm simply initiates from the point with the highest response, discards feature points within its 8-neighbor vicinity, and then sequentially examines the remaining points, in this article, steps c and d are employed to ensure that the selected feature points are uniformly distributed across the entire image. Steps c1 to c2 denote the search range estimation (depicted as red boxes) through a binary search process, which iterates until the number of global points is achieved. Step d represents the iterative adjustment of the search range estimation (illustrated with red boxes) to ultimately achieve the required number of feature points, wherein the formula for the search range estimation is $d_l = \frac{1}{2}\sqrt{\frac{n}{m}}$ and $d_h = \frac{H+W+2m-\sqrt{\Delta}}{2(m-1)}$ (where m is the required quantity of keypoints after homogenization, n is the initial value of keypoints, H and W are the height and width of the image, respectively, and $\Delta = 4w + 4m + 4Hm + H^2 + w^2 - 2WH + 4WHm$). The c process is repeated until the number of homogenized keypoints becomes equal to the required number of keypoints.

C. Multiscale Feature Descriptor Construction

Keypoint descriptor constructed from a single-scale support region cannot achieve good results for SAR image keypoint matching. To increase the accuracy of feature matching, in [5] and [20], multiscale spatial information was used to construct descriptors. This approach was aimed at solving the problem that two noncorresponding points may have similar feature information with a keypoint descriptor constructed in single-scale space in case of noisy images. Based on the construction of multiscale descriptors in [5], this study uses more scale space information to construct feature descriptors to further improve their uniqueness and increase the accuracy of feature matching.

In this study, four concentric circle support regions are used to build a feature GLOH descriptor. The four different regions generated by the four concentric circles record different scale space information. The scale space information is calculated using (7). As depicted in Fig. 4(a), the SAR–SIFT method utilizes the GLOH approach for the construction of descriptors, where the radii of the outer, middle, and inner circles are in the ratio of 1:0.73:0.25. Fig. 4(b) illustrates that PSO-SIFT builds upon the SAR–SIFT descriptors, dividing the SAR–SIFT descriptor into 17 positional bins according to a 45° rule, thereby enhancing the distinctiveness of the descriptor. Considering that a single-scale support region may occasionally lead to mismatched feature

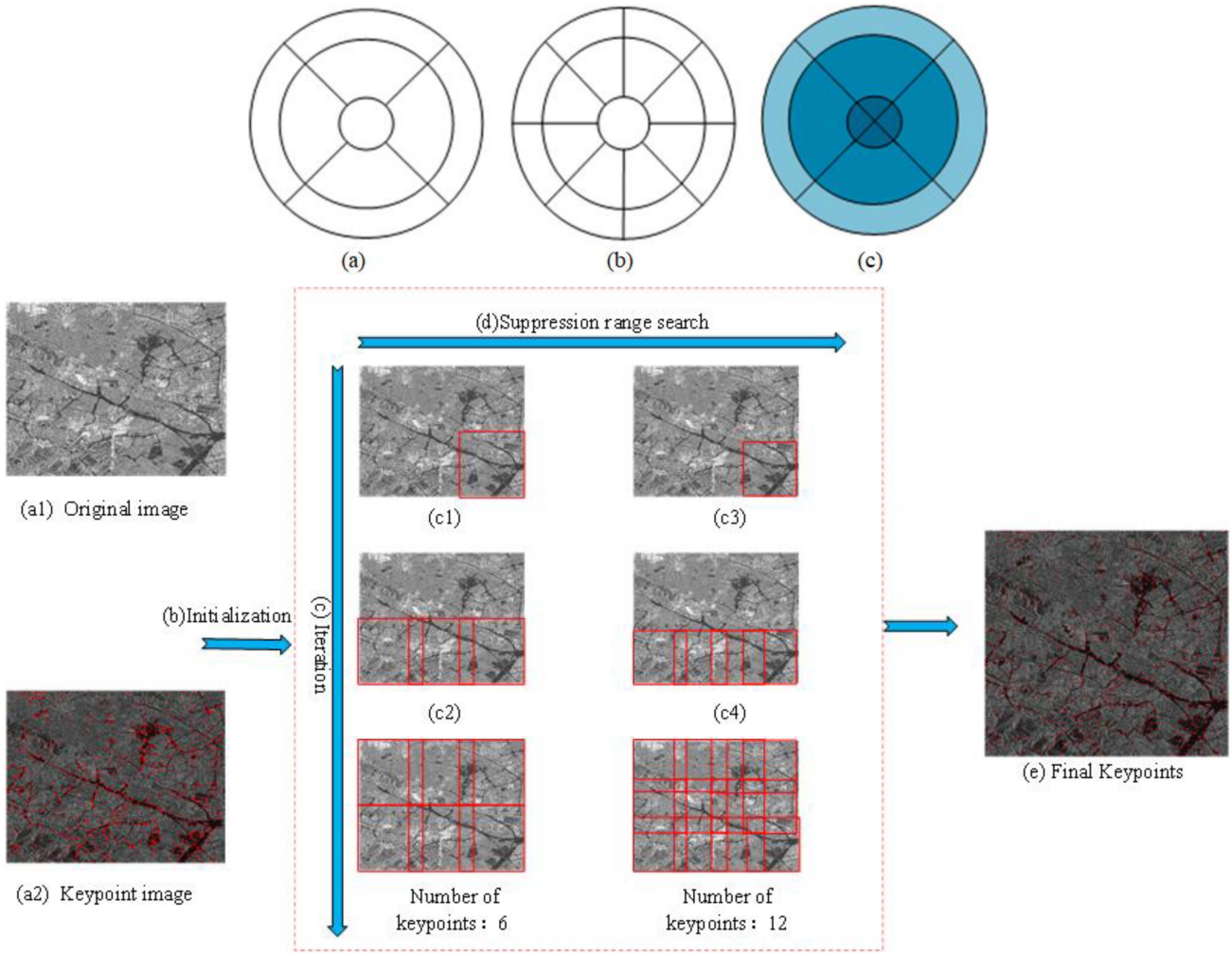


Fig. 4. Descriptor structure diagram. (a) Multiscale spatial information descriptor construction map. (b) Descriptor internal division.

TABLE I
DESCRIPTOR EXPERIMENT

Ratio	RMSE	CMR
1:0.73:0.7:0.25	1.13	70.1%
1:0.73:0.65:0.25	0.92	85.6%
1:0.73:0.6:0.25	0.86	86.8%
1:0.73:0.5:0.25	0.78	91.3%
1:0.73:0.45:0.25	0.83	88.4%
1:0.73:0.4:0.25	0.93	85.6%
1:0.73:0.35:0.25	1.21	79.3%
1:0.73:0.3:0.25	1.34	71.7%

points exhibiting similar local features, as shown in Fig. 4(c), I-SAR-SIFT employs multiple concentric circular support regions at different scales to address this issue. However, this method still divides the positional regions at 90°, which can result in decreased matching accuracy when matching similar terrain areas in the image. Fig. 4(d) demonstrates that the descriptor construction in this article combines the aforementioned methods, adopting both the 45° division scheme to reduce matching accuracy in similar terrain regions and multiscale concentric circular regions to further enhance the uniqueness of the descriptor. This article continues to use the ratio of concentric circle radii as adopted in SAR–SIFT. For the newly added concentric circle areas, this study has determined the optimal radius ratio through experiments. The experimental results are shown in Table I. The experiment utilized the Test-2 data provided in Table III for the study. It can be observed that the ratios exhibit an extremum

TABLE II
TEST ENVIRONMENT

Hardware environment	Software environment
CPU: AMD-5800 X	Operating system: Windows10 x64
Graphics: RTX1650	Visual Studio 2017
Memory: 16 GB	Opencv 3.4.5

when set at 1:0.73:0.5:0.25 for both quantitative indicators. Consequently, this article selects this ratio as the ratio for the GLOH descriptor.

D. Matching and Outlier Detection

After the keypoints are obtained, an outlier elimination method combining the nearest neighbor distance ratio (NNDR) and the random sample consensus (RANSAC) [30] is used for feature matching. Based on the Euclidean distance, the NNDR selects the closest distance and the next closest distance by sorting. If a keypoint satisfies

$$d_n < d_{sn} \times d_{\text{ratio}} \tag{8}$$

it is considered as a preliminary matching point. In (8), d_n represents the nearest neighbor distance, d_{sn} represents the second nearest neighbor distance, and d_{ratio} represents the matching threshold. In this study, $d_{\text{ratio}} = 0.8$.

III. EXPERIMENTS AND ANALYSIS

In this part, we propose the testing and analysis of the method issued in this article conducted on five sets of SAR image data (four different satellites are selected). The images are compared, and the final keypoint matching results on the five sets are tested and analyzed using three quantitative evaluation criteria: RMSE, keypoint repeatability, correct matching ratio (CMR), number of matches (NCM), mutual information (MI), and distribution quality. The specific test environment is detailed in Table II, and the quantitative metrics are as follows:

A. Transformation Model

1) RMSE: To facilitate a more quantitative assessment, it is necessary to obtain the ground truth geometric transformation between each pair of images. However, due to various interferences, real datasets often lack true geometric transformations, and approximate ground truth geometric transformations are commonly used for evaluation. The determination of an affine transformation matrix requires at least three pairs of corresponding points. Li et al. [32] utilize five pairs of points that are evenly distributed and essentially cover the entire global area to solve for the affine transformation matrix. The test image selected in this article is slightly larger than the test image in the above method, so this study selects ten pairs of points that are evenly distributed and essentially cover the entire global area to solve for the affine transformation matrix, serving as an approximation of the ground truth. Moreover, the residual was calculated according to the estimated affine transformation model, and the RMSE was defined as follows:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[(x_i - x_i{}')^2 + \left(y_i - y_i'\right)^2\right]} \\ (x_i', y_i', 1) &= H \cdot (x_i, y_i, 1)\,. \end{aligned} \tag{9}$$

The parameter N represents all the keypoints extracted by the feature extraction algorithm, and (x_i, y_i) represents the keypoints obtained after RANSAC. (x_i', y_i') represents the location of (x_i, y_i) after being transformed by the estimated affine transformation model.

2) CMR: The CMR is defined as $\text{CMR} = \frac{N_{\text{corr}}}{N_{\text{nndr}}}$, where N_{corr} depends on manually selecting ten uniformly distributed matching point pairs, to evaluate the affine model for the image pairs of each test. We utilize this affine model to determine the location errors of the matches received, when the matching point obtained by the matching algorithm is within 2 pixels of the point obtained by the homography matrix transformation, the matching point is considered to be the correct point. N_{nndr} represents an initial matching point pair that is not screened by RANSAC.

3) Repeatability: It is the ratio of the correct pairing rate to the minimum number of features detected in two images. The repetition rate is defined as follows:

$$R = \frac{N_{\text{repeat}}}{\min(N_1, N_2)} = \frac{|\{\| x_i - Hx_i'\| < 3\}_i^{N1}|}{\min(N_1, N_2)} \tag{10}$$

where N_{repeat} represents the number of duplicate points in the two images, N_1, N_2 express the number of keypoints initially detected in the two images, and H is the estimated ground truth transformation between the two images.

4) NCM: Li et al. [32] employed the NCM as an evaluation metric for the number of matched points. In this article, we continue to use this metric as the basis for evaluating the quantity of matched points. The NCM is determined by the following strategy: The affine transformation matrix obtained through (9) is used to calculate the residual value for each pair of corresponding points. Pairs with residuals less than two pixels are considered to be correctly matched.

5) MI: MI method utilized by Nabatchian et al. [33] is adept at assessing the similarity of matched points. Hence, this article continues to employ this method for evaluating the similarity of matched points. The MI calculation process is as follows: Based on the coordinates of the feature points in the left and right images, a 4x4 neighborhood is constructed around each pair of matching points. The grayscale values within these neighborhoods, which range from 0 to 255, are considered as two variables. The information entropy of the feature points is calculated using (11). Subsequently, the joint entropy of the matching points is determined using (13). Finally, the mutual information of the matching points is computed using (12),with detailed computational specifics available in [34]

$$\text{MI}(A, B) = H(A) + H(B) - H(A, B) \tag{11}$$

$$H(\mathrm{x}) = -\sum_{\mathrm{x}\in X} p(\mathrm{x}) \log(p(\mathrm{x})) \tag{12}$$

TABLE III
DETAILED INFORMATION OF ALL TEST CASES

Test	Image	Size	Data	Location
Test-1	Gaofen-3 satellite, 3 m	1437 × 1333	24/10/2018	Shanghai, China
	Gaofen-3 satellite, 3 m	1251 × 1330	02/01/2019	Shanghai, China
Test-2	Gaofen-3 satellite, 3 m	1437 × 1332	21/04/2018	Shanghai, China
	Gaofen-3 satellite, 3 m	1268 × 1332	24/10/2018	Shanghai, China
Test-3	RADARSAT, 8 m	600 × 500	2/06/2008	Yellow river, China
	RADARSAT, 8 m	600 × 500	2/06/2009	Yellow river, China
Test-4	Sentinel-1A, 10 m	1142 × 874	23/12/19	Myanmar
	Sentinel-1A, 10 m	1713 × 1293	11/12/19	Myanmar

$$H(A,B) = -\sum_{x,y} p(x,y)\log\left(p(x,y)\right) \tag{13}$$

where MI represents mutual information value, *H(A)* represents information entropy of Figure *A*, *H(B)* represents information entropy of Figure *B*, *H(A, B)* represents joint entropy, *x* represents gray value, *p*(.) represents the probability that the grayscale value is *x* in grayscale information.

6) Local and Global Distributions of Matches: The distribution quality of the matched point pairs is assessed using two well-known measures. The first is the criterion introduced in [34]. The area descriptor is a measure of the dispersion or variability of the triangular areas formed by the matching points. A smaller area descriptor indicates a better distribution of matching points. The calculation method is illustrated in (14), with detailed computational specifics provided in [34]. The second evaluation metric is the comprehensive metric ϕ, as adopted by Zhu et al. [35]. As the second measure, the statistical distribution quality factor ϕ, developed in paper [35], a smaller ϕ indicates a better distribution of matching points. The calculation method for this metric is shown in (15), with detailed computational specifics available in [36]

$$D = D_A \times D_S = \sqrt{\frac{\sum_{i=1}^{n}\left(\frac{A_i}{\bar{A}}-1\right)}{n-1}} \times \sqrt{\frac{\sum_{i=1}^{n}\left(S_i-1\right)}{n-1}},$$
$$\bar{A} = \frac{\sum_{i=1}^{n} A_i}{n};\ S_i = \frac{3\max\left(J_i\right)}{\pi}. \tag{14}$$

$$\phi = \frac{2\times\left(1/N_{\text{red}} + \text{rms}_{\text{LOO}} + \text{BPP}\left(1.0\right) + S_{\text{cat}}\right) + \text{rms}_{\text{all}} + 1.5\times\left(p_{\text{quad}} + S_{\text{kew}}\right)}{(2+2+2+2+1+1.5)}. \tag{15}$$

In this equation, *n* is the total number of Delaunay triangles, Ai and $\max(J_i)$ indicate the area and maximum angle (in radians) of the *i*th triangle, and *A* is the mean area of all triangles. Lower *D* values show the geometrical homogeneity of the obtained Delaunay triangles, which is a measure of the uniform distribution of the matched points. N_{red} refers to the number of matching pairs, rms_{all} refers to the root-mean-square error for detected matching pairs, rms_{LOO} is a measure combining the leave one-out method with rms_{all}, BPP(1.0) means the percentage of bad points whose residual distances greater than 1 pixels, p_{quad} is used to detect whether the reserved keypoints are equally distributed across the quadrants, S_{kew} is a statistical evaluation regarding the presence of a preference axis on the residual scatter plot, and S_{cat} is a statistical evaluation regarding the feature point distribution across the whole image.

B. Data Sources and Parameter Settings

The data used in this study are images of different regions acquired at different times from four different satellites. The satellite image data of these different times show significant radiative changes and geometric differences, and the image details can be found in Table III. We conduct experiments using different feature extraction algorithms on five sets of image data. The parameter settings for the SAR–SIFT algorithm are as follows: To limit the computational load and prevent excessive consumption of computational resources, this article restricts the number of image keypoints to 4000. Following the scale-space parameter recommendations provided by the SIFT algorithm, the first layer of the scale space in this article is set to $\rho_1 = 2$, the constants of two adjacent scale spaces are $\sqrt[3]{2}$, the number of scale spaces is set as 8, and arbitrary parameter *d* is set as 0.04. The parameter settings of the UDR detector are adjusted based on the image. The selection of UDR parameters is based on the parameter setting requirements proposed in the literature [28], with specific parameter adjustment requirements as described in the following formula:

$$\begin{cases} \alpha = \{2,3,4,5\} \\ \beta = 0.5 + 0.5k_1, & k_1 = 0,1,\ldots,9 \\ l_{\|} = 1 + k_2, & k_2 = 0,1,\ldots,4 \\ \sigma_{\|} = 1.5 + 0.5k_3, & k_3 = 0,1,\ldots,13. \end{cases} \tag{16}$$

In SAR–SIFT, a threshold of 0.8 is adopted for the NNDR method. This is an empirically determined value that is considered to yield good results. It balances false matches and missed matches, providing reasonable matching quality. Therefore, this article adheres to the threshold selection used in SAR–SIFT.

C. Comparison of Results of the Proposed Algorithm

In this section, the convolutional, multiscale gradient magnitude, and Harris–Laplace scale space images obtained by applying the UDR detector and the ROEWA operator are shown. Based on the qualitative analysis of the advantages of the UDR detector compared to the ROEWA operator, the method proposed in this article and the SAR–SIFT algorithm respectively extract keypoints from simulated SAR images. Accordingly, the difference in the distributions of the keypoints is obtained.

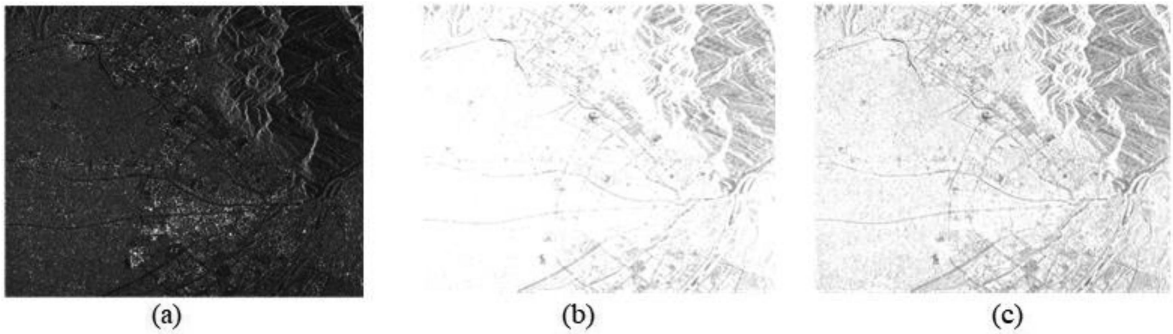


Fig. 5. Convolutional image comparison. (a) Original image. (b) Convolutional image obtained with ROEWA detector. (c) Convolutional image obtained with UDR detector.

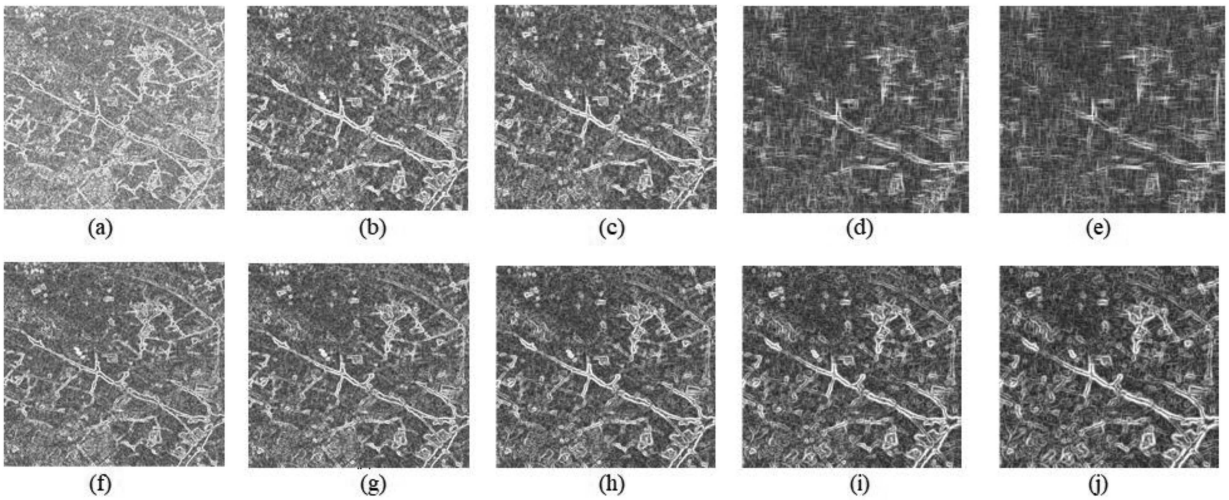


Fig. 6. Gradient images at different scale spaces. (a)–(e) Gradient images obtained by ROEWA operator at different scales and (f)–(j) gradient images obtained by UDR detector at different scales. The scale values from left to right are 4.00, 5.04, 6.35, 8.00, and 10.08, respectively.

1) Convolving Images With Different Detectors

Fig. 5 shows the convolutional images obtained with different operators in the same scale space. After the ROEWA operator convolves the original image, the image is subject to a large noise, and good texture information cannot be derived from the entire image. However, the convolutional image after application of the UDR detector to the original image is highly refined.

2) Image Gradients Obtained By Different Detectors: Fig. 6 presents the image gradients obtained using the GR computation method with different detectors across various spatial scales. At lower scales, compared to the SAR-SIFT algorithm, our algorithm captures image gradients that retain texture information more closely and preserve edge information of the image effectively. As the spatial scale increases, the image gradient information preserved by both our algorithm and the SAR-SIFT algorithm begins to blur gradually. However, at higher scales, the image gradients retained by the SAR-SIFT images are less clear and contain less rich texture information compared to our algorithm.

3) Multiscale Harris–Laplace Images Constructed By Different Detectors: Fig. 7 shows the Harris–Laplace scale space images constructed by applying the ROEWA and UDR detector in different scale spaces on SAR images. In this study, a five-layer scale space was built for the experiments. With the increase in the scale factor, scale images can be obtained, and the feature information gradually decreases. Compared to the Harris–Laplace scale space image constructed by the ROEWA operator, those constructed by the UDR detector have fewer bright patches; therefore, the obtained image information is also clearer.

4) Keypoint Detection in Simulated Images By Different Detectors: Fig. 8 shows the results of two different corner detection methods are given. The one on the left is the corner points obtained by Harris–Laplace–Roewa. Although there are many corner points, they are not evenly distributed in or around the rectangle, and some of them deviate from the corner points. The one on the right is the corner points obtained by the method in this article, although there are fewer corner points compared with the one on the left, they are concentrated in the four corner points. In contrast, the detector of our method performs better.

5) Keypoint Detection in Local and Global Image: Fig. 9 shows the distribution of local keypoints. The left image illustrates the local distribution of keypoints obtained by SAR-SIFT. It can be clearly observed that there is a bias in the localization of keypoints in the lower left and lower right areas. In contrast, the proposed algorithm can accurately locate keypoints on the edges of the image.

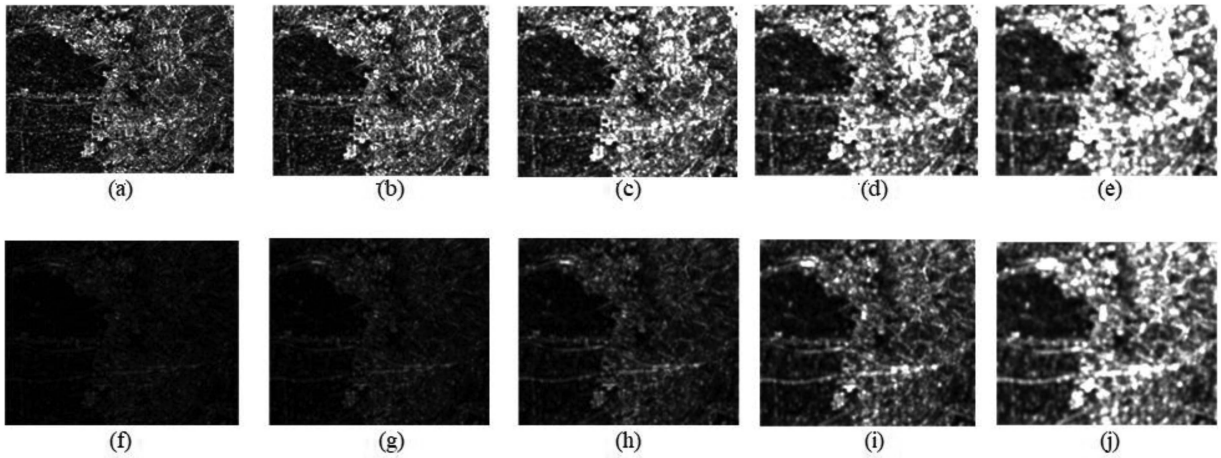


Fig. 7. Gradient Harris–Laplace scale space images by different detectors. (a)–(e) Harris–Laplace scale space images by ROEWA operator. (f)–(j) Harris–Laplace scale space images by UDR detector. The scale values from left to right are 4.00, 5.04, 6.35, 8.00, and 10.08, respectively.

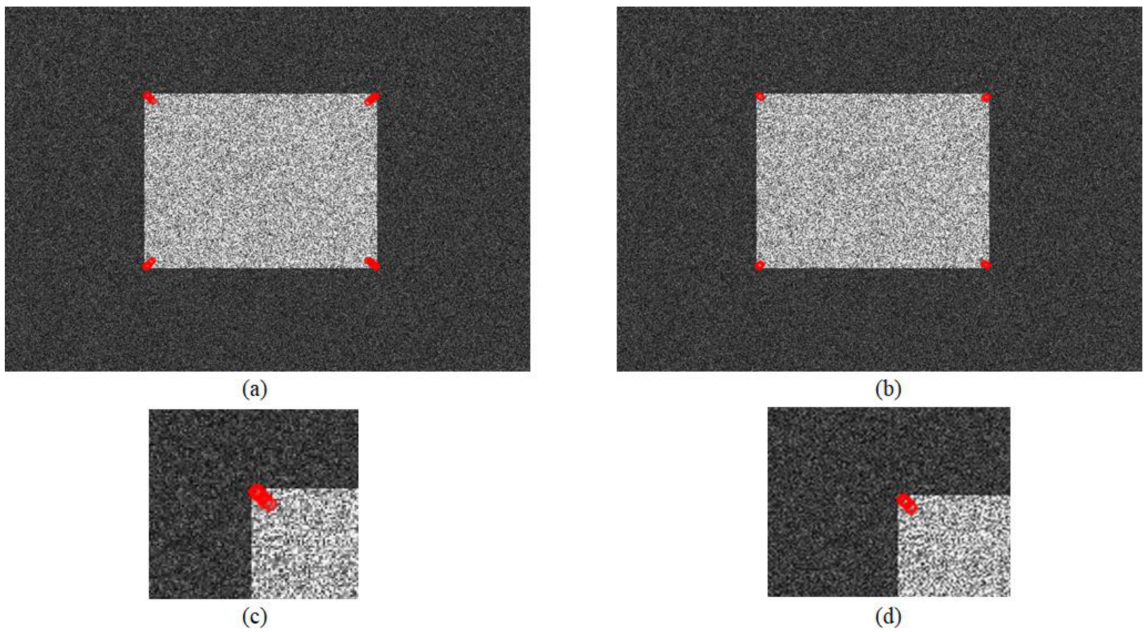


Fig. 8. Corner detection in simulated SAR image. (a) Harris–Laplace ROEWA operator. (b) Harris–Laplace UDR detector. (c) Harris–Laplace ROEWA operator and enlarged corner. (d) Harris–Laplace UDR detector and enlarged corner.

6) Feature Point Screening Results and Efficient Comparison Results: Upon examination of Fig. 10, it is evident that there is a significant concentration of closely overlapping feature points within certain regions of the image. This high density of feature points not only results in redundant information but also poses a high risk of erroneous matches during subsequent feature matching processes due to their high degree of similarity. After processing with the algorithm presented in this article, the distribution of feature points has been significantly optimized. In areas of high density, the algorithm effectively prevents excessive suppression, ensuring a suitable retention of feature points. Conversely, in regions of low density, the algorithm preserves a greater number of feature points to maintain the integrity and richness of the feature point distribution. This approach not only mitigates the clustering of feature points in local areas but also enhances their distinctiveness.

7) Experimental Results and Evaluation: To verify that the SAR image matching framework proposed in this article has better feature matching accuracy, we use multiple sets of different satellite image data to conduct comparative experiments with SIFT, SAR–SIFT, speeded up robust features (SURF) algorithms, Affine-SIFT (ASIFT) and KAZE-SAR [31], OS-SIFT [23], and EEKHI–SAR–SIFT retained 40% of the original keypoints using SSC. The matching consequences of the different algorithms on the test data (Gaofen-3, RADARSAT, and Sentinel-1A) are shown in Figs. 11–14, where the red line represents the final correct matching result, and the green line represents the wrong matching result (based on visual

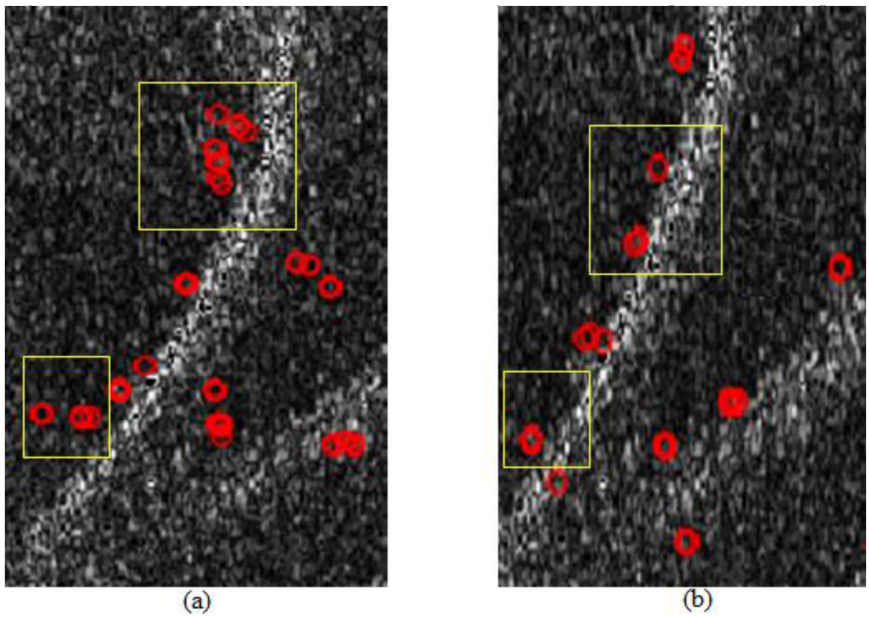


Fig. 9. Position of keypoints of SAR–SIFT and the method in this article. (a) SAR—SIFT. (b) This article.

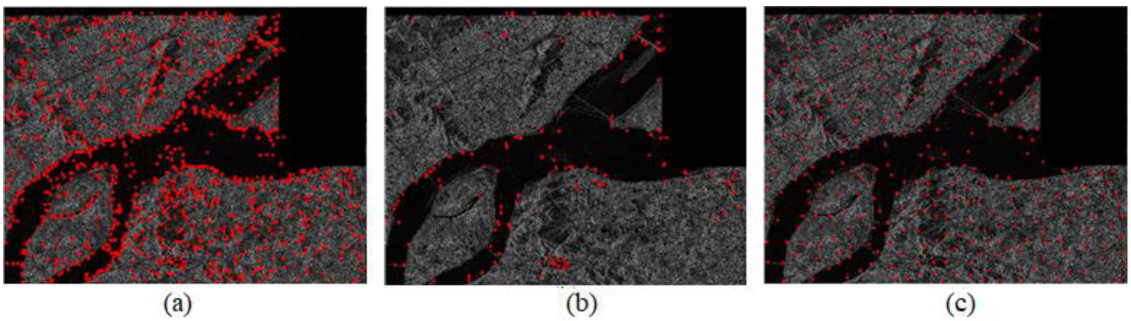


Fig. 10. Feature point distribution comparison diagram. (a) Initial feature point. (b) TopM. (c) This article.

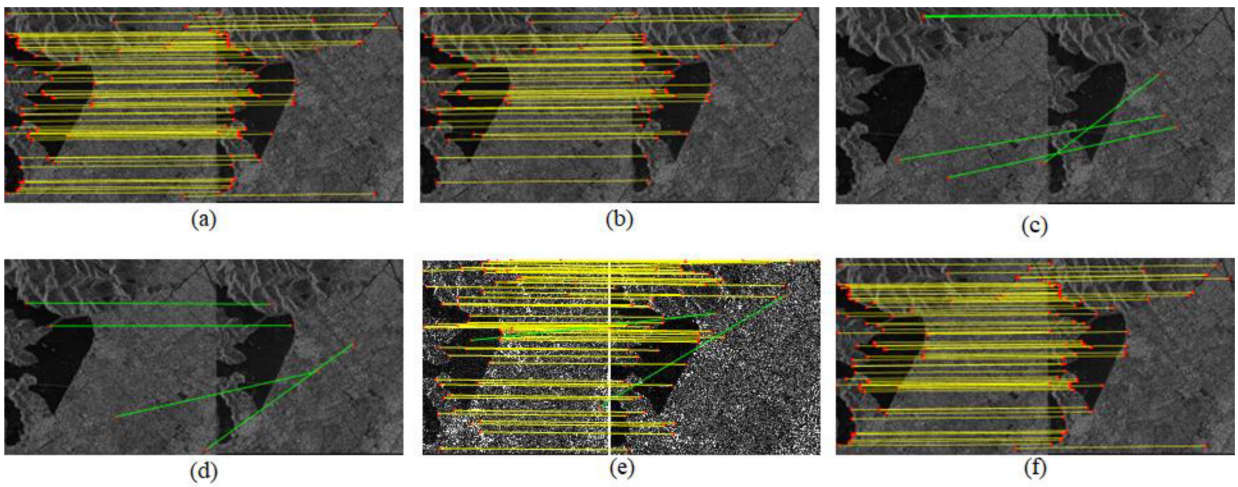


Fig. 11. Matching points obtained by methods used in experiments for Test-1. (a) EEKHI-SAR-SIFT, (b) SARSIFT, (c) SIFT, (d) SURF, (e) KAZE-SAR, (f) OS-SIFT.

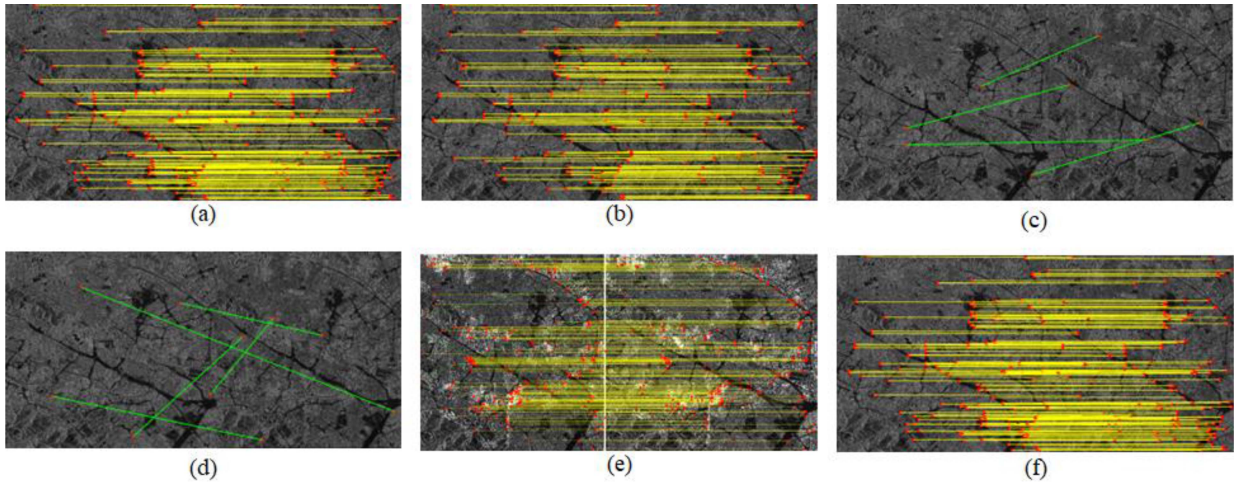


Fig. 12. Matching points obtained by methods used in experiments for Test-2. (a) EEKHI-SAR-SIFT, (b) SARSIFT, (c) SIFT, (d) SURF, (e) KAZE-SAR, (f) OS-SIFT.

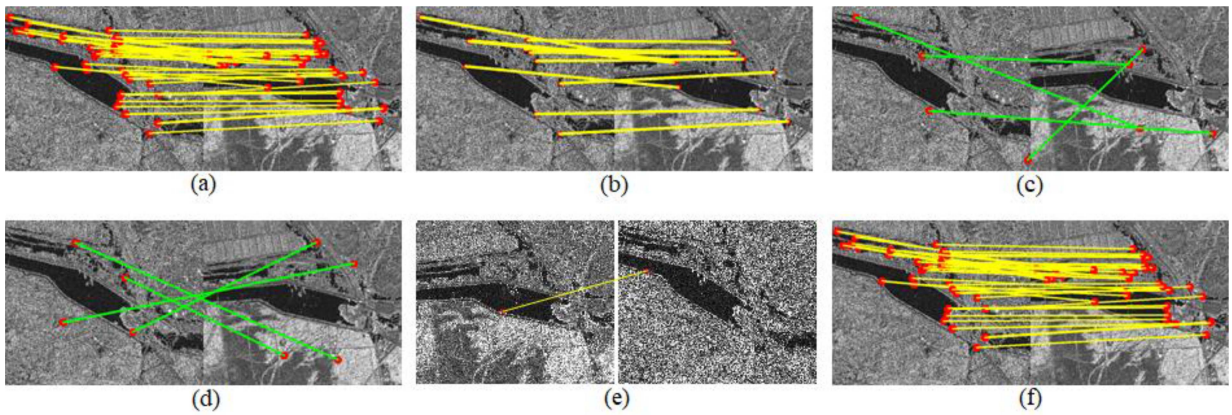


Fig. 13. Matching points obtained by methods used in experiments for Test-3. (a) EEKHI-SAR-SIFT, (b) SARSIFT, (c) SIFT, (d) SURF, (e) KAZE-SAR, (f) OS-SIFT.

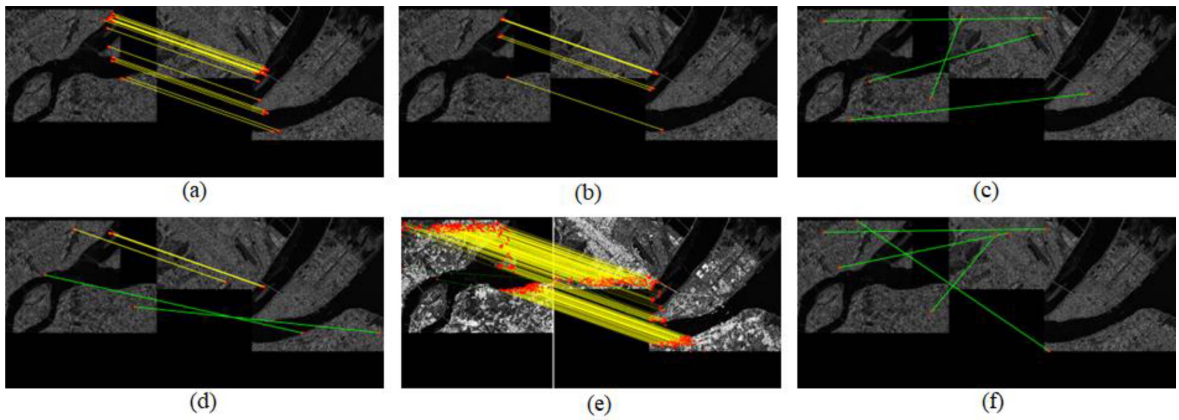


Fig. 14. Matching points obtained by methods used in experiments for Test-4. (a) EEKHI-SAR-SIFT, (b) SARSIFT, (c) SIFT, (d) SURF, (e) KAZE-SAR, (f) OS-SIFT.

TABLE IV
QUANTITATIVE COMPARISON OF MATCHING RESULTS TEST

Test	Method	RMSE	Repeatability	CMR	NCM	MI	D	ϕ
1	SAR–SIFT	1.714	0.016	68.5%	113	0.651	0.683	0.781
	EEKHI–SAR–SIFT	**0.534**	**0.084**	**84.2%**	**118**	**0.893**	**0.452**	**0.515**
	SIFT	760.261	*	0	0	*	*	*
	SURF	901.400	*	0	0	*	*	*
	KAZE-SAR	1.823	0.058	72.9%	56	0.737	0.552	0.644
	OS-SIFT	1.934	0.031	69.4%	64	0.644	0.564	0.712
2	SAR–SIFT	3.273	0.019	67.4%	103	0.697	0.704	0.765
	EEKHI–SAR–SIFT	**0.785**	**0.045**	**81.7%**	**134**	**0.843**	**0.464**	**0.536**
	SIFT	472.169	0.024	0	0	*	*	*
	SURF	6.136	0.009	0	0	*	*	*
	KAZE-SAR	2.309	0.391	71.5%	116	0.628	0.589	0.631
	OS-SIFT	2.435	0.243	67.8%	108	0.586	0.632	0.667
3	SAR–SIFT	1.811	0.038	51.2%	17	0.545	1.124	1.164
	EEKHI–SAR–SIFT	**1.424**	**0.089**	**76.4%**	**21**	**0.817**	**0.735**	**0.617**
	SIFT	12.884	0.045	0	0	*	*	*
	SURF	9.3607	0.056	0	0	*	*	*
	KAZE-SAR	1.582	0.073	0	0	*	*	*
	OS-SIFT	1.634	0.072	65.3%	12	0.587	0.943	0.678
4	SAR–SIFT	1.286	0.036	42.6%	12	0.513	1.624	1.534
	EEKHI–SAR–SIFT	**0.7246**	**0.092**	**67.9%**	**21**	**0.792**	**0.835**	**0.794**
	SIFT	820.957	0	0	0	0	*	*
	SURF	106.394	0	0	0	0	*	*
	KAZE-SAR	3.189	0.027	60.3%	18	0.536	1.153	0.801
	OS-SIFT	57.543	0	0	0	*	*	*

*Note that due to the limited number of matching points, it is not possible to obtain accurate results.
The bold values represent the best value in each metric.

judgment). The quantitative evaluation results are presented in Table IV, the algorithm presented in this article has achieved the best quantitative metrics in the aforementioned evaluation criteria. MI quantitative results of our algorithm are consistently around 0.8, indicating that the matched points obtained by our algorithm exhibit a good degree of similarity. In contrast, the MI results of other algorithms are inferior to those of our algorithm, and they are unable to provide quantitative analysis results for all test images, demonstrating a poorer general applicability. The D and ϕ indices of our algorithm, as observed in the Test-1 and Test-2 experimental data, are maintained within the range of 0.45 to 0.55. Due to the lower image quality in Test-3 and Test-4, the quantitative assessment results are not as favorable as those in Test-1 and Test-2. Nonetheless, even in these conditions, the distribution of matched points by our algorithm still outperforms other algorithms.

Fig. 11 presents the final matched point pairs obtained using GF-3 imagery, which has regions with rich texture information. Compared to other algorithms, Fig. 11(a) exhibits a more favorable distribution of matched points and lacks obvious erroneous matches. As shown in Table IV, Fig. 11(a) achieves the highest accuracy. In contrast, although Fig. 11(b) appears to have no apparent errors, Table IV indicates lower accuracy compared to the methods in Fig. 11(a) and (e). This limitation arises from the poor noise-filtering capability of its own filter, and the uneven distribution of matched points in Fig. 11(b) with no matches at the top may lead to reduced accuracy in subsequent adjustment or 3-D reconstruction, which is critically detrimental. Fig. 11(c) and (d) fail to produce correct results due to limitations in their respective algorithms, preventing effective handling of highly noisy imagery. Fig. 11(e) yields more matched points but also contains noticeable erroneous matches. This is because the method employs nonlinear diffusion filtering to reduce noise influence, which can mitigate noise to a certain extent. However, due to the similarity between speckle noise and image details, the remaining unfiltered noise significantly affects the final results. Moreover, the loss of image details and blurring caused by the smoothing process of nonlinear diffusion filtering negatively impact the accuracy. Quantitative evaluations in Table III reveal a slight disparity in accuracy between Fig. 11(e) and (a). Although Fig. 11(f) lacks obvious erroneous points, the method improves on scale space and corner points. However, when corner localization is imprecise, it fails to achieve higher accuracy, as reflected in the lower quantitative evaluation metrics in Table IV. The situations depicted in Figs. 12 and 13 resemble Fig. 11, where SIFT and SURF methods in Fig. 13 do not exhibit apparent errors but still show significant positional deviations upon closer inspection, as demonstrated by the quantitative metrics in Table IV.

Fig. 14 illustrates the results obtained from Sentinel data, characterized by poor image quality and high noise levels. The actual overlapping area is only a small portion on the right side, contributing to the overall lower accuracy of all methods. Fig. 14(f) fails to produce correct results despite utilizing the stability of corner points and modifying the DOG scale space of SIFT to Harris scale space. This modification does not effectively address noise influence, leading to incorrect results when corner points are severely disrupted by noise. Additionally, due to the small overlapping area in the Sentinel imagery, numerous useless feature points are generated, further affecting the matching process. Consequently, Fig. 14(f) does not yield any correct matches. Fig. 14(e) contains 2–3 obvious erroneous

matches, similar to the issues observed in Figs. 12–14. Fig. 14(c) and (d) do not yield correct results due to their limited noise resistance capability. Although no apparent erroneous points are observed in Fig. 14(b), the method's weak edge localization ability results in lower accuracy of the final results. The method in Fig. 14(a) achieves the highest accuracy as its results exhibit a uniform distribution. Furthermore, it demonstrates superior noise resistance and preserves image details, as depicted in Table IV's quantitative evaluation results.

IV. Conclusion

A novel SAR image keypoint extraction algorithm based on SAR–SIFT, called EEKHI–SAR–SIFT, is developed, which improves the keypoint extraction accuracy of SAR images and yields uniformly distributed matching point pairs. With a UDR edge detector used, more accurate image edge information is obtained than the ROEWA detector, and subsequently a non-maximum suppression algorithm is employed to homogenize the keypoints. Thus, we can consider the quantity of keypoints and the repetition rate simultaneously, and a multiscale descriptor construction method is used to increase the distinctiveness of the feature descriptor. The outcomes of experiment indicate the repetition rate of the keypoints extracted by the EEKHI–SAR–SIFT method is increased by approximately 5% compared to that by the SAR–SIFT algorithm. Moreover, the RMSE is decreased by approximately 1 pixel, whereas the quantity of matching points is not significantly decreased with the above improvements. Tests using multiple sets of SAR image data from different satellites (Gaofen-3, RADARSAT, and Sentinel-1A) prove the universality of the EEKHI–SAR–SIFT method. The performance of the UDR edge detector employed in this article is highly sensitive to parameter selection. Improper parameter settings may adversely affect the filtering outcomes. Therefore, to achieve optimal results, it is necessary to adjust parameters according to the specific characteristics of different image data. However, this requirement limits the practical applicability of the method. Designing adaptive parameter adjustment strategies will be the main focus of our subsequent research efforts.

References

[1] F. Song et al., "Multi-scale feature based land cover change detection in mountainous terrain using multi-temporal and multi-sensor remote sensing images," *IEEE Access*, vol. 6, pp. 77494–77508, 2018, doi: 10.1109/access.2018.2883254.

[2] Q. Yu, S. Zhou, Y. Jiang, P. Wu, and Y. Xu, "High-performance SAR image matching using improved SIFT framework based on rolling guidance filter and ROEWA-powered feature," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 12, no. 3, pp. 920–933, Mar. 2019, doi: 10.1109/jstars.2019.2897171.

[3] X. Li, J. Yang, and J. Ma, "Large scale category-structured image retrieval for object identification through supervised learning of CNN and SURF-based matching," *IEEE Access*, vol. 8, pp. 57796–57809, 2020, doi: 10.1109/access.2020.2982560.

[4] S. Chen, S. Zhong, B. Xue, X. Li, L. Zhao, and C. -I. Chang, "Iterative scale-invariant feature transform for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3244–3265, Apr. 2021, doi: 10.1109/tgrs.2020.3008609.

[5] S. Paul and U. C. Pati, "SAR image registration using an improved SAR-SIFT algorithm and delaunay-triangulation-based local matching," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 12, no. 8, pp. 2958–2966, Aug. 2019, doi: 10.1109/jstars.2019.2918211.

[6] W. Ouyang, T. Zhao, and W.-K. Cham, and L. Wei, "Fast full-search-equivalent pattern matching using asymmetric haar wavelet packets," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 4, pp. 819–833, Apr. 2018, doi: 10.1109/tcsvt.2016.2629621.

[7] J. Li, W. Xu, P. Shi, Y. Zhang, and Q. Hu, "LNIFT: Locally normalized image for rotation invariant multimodal feature matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5621314, doi: 10.1109/tgrs.2022.3165940.

[8] X. Xu et al., "Multimodal registration of remotely sensed images based on Jeffrey's divergence," *ISPRS J. Photogramm. Remote Sens.*, vol. 122, pp. 97–115, 2016, doi: 10.1016/j.isprsjprs.2016.10.005.

[9] F. Zhao, Q. Huang, and W. Gao, "Image matching by normalized cross-correlation," in *Proc. IEEE Int. Conf. Acoust.*, Jul. 2006, pp. II–II, doi: 10.1109/ICASSP.2006.1660446.

[10] Y. Hel-Or, H. Hel-Or, and E. David, "Fast template matching in non-linear tone-mapped images," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 1355–1362, doi: 10.1109/ICCV.2011.6126389.

[11] D. I. Barnea, "A class of algorithms for fast digital image registration," *IEEE Trans.Comput.*, vol. C-21, no. 2, pp. 179–186, Feb. 1972, doi: 10.1109/TC.1972.5008923.

[12] A. A. Cole-Rhodes, K. L. Johnson, J. Lemoigne, and I. Zavorin, "Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1495–1511, Dec. 2003, doi: 10.1109/TIP.2003.819237.

[13] P. C. Ng and S. Henikoff, "SIFT: Predicting amino acid changes that affect protein function," *Nucleic Acids Res.*, vol. 31, no. 13, pp. 3812–3814, Jul. 2003, doi: 10.1093/nar/gkg509.

[14] L. Meng et al., "Investigation and evaluation of algorithms for unmanned aerial vehicle multispectral image registration," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 102, 2021, Art. no. 102403, doi: 10.1016/j.jag.2021.102403.

[15] S. Castillo-Carrión and J.-E. Guerrero-Ginel, "SIFT optimization and automation for matching images from multiple temporal sources," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 57, pp. 113–122, 2017, doi: 10.1016/j.jag.2016.12.017.

[16] H. Kartal, U. Alganci, and E. Sertel, "Automated orthorectification of VHR satellite images by SIFT-based RPC refinement," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 6, 2018, doi: 10.3390/ijgi7060229.

[17] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 6, pp. 764–773, Nov. 1988, doi: 10.1109/36.7708.

[18] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005, doi: 10.1109/TPAMI.2005.188.

[19] W. Bangsong, Z. Jixian, L. Lijun, H. Guoman, and Z. Zhao, "A uniform SIFT-like algorithm for SAR image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 7, pp. 1426–1430, Jul. 2015, doi: 10.1109/lgrs.2015.2406336.

[20] S. Paul and U. C. Pati, "A Gabor odd filter-based ratio operator for SAR image matching," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 3, pp. 397–401, Mar. 2019, doi: 10.1109/lgrs.2018.2872979.

[21] C. Han et al., "An image matching method for SAR orthophotos from adjacent orbits in large area based on SAR-Moravec," *Remote Sens.*, vol. 12, no. 18, 2020, Art. no. 2892, doi: 10.3390/rs12182892.

[22] J. Fan, Y. Wu, M. Li, W. Liang, W. Liang, and Q. Zhang, "SAR image registration using multiscale image patch features with sparse representation," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 10, no. 4, pp. 1483–1493, Apr. 2017, doi: 10.1109/JSTARS.2016.2628911.

[23] Y. Xiang, F. Wang, and H. You, "OS-SIFT: A robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3078–3090, Jun. 2018, doi: 10.1109/tgrs.2018.2790483.

[24] Y. Hong et al., "SAR image registration based on ROEWA-blocks and multiscale circle descriptor," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 10614–10627, 2021, doi: 10.1109/jstars.2021.3119923.

[25] Y. Xiang, L. Peng, F. Wang, and X. Qiu, "Fast registration of Multiview slant-range SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2021, Art. no. 4007505, doi: 10.1109/lgrs.2020.3045099.

[26] D. Xiang, Y. Xu, J. Cheng, Y. Xie, and D. Guan, "Progressive keypoint detection with dense Siamese network for SAR image registration," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 5, pp. 5847–5858, Oct. 2023, doi: 10.1109/TAES.2023.3266415.

[27] D. Xiang, H. Ding, X. Sun, J. Cheng, C. Hu, and Y. Su, "PolSAR image registration combining Siamese multiscale attention network and joint filter," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, Mar. 2024, Art. no. 5208414, doi: 10.1109/TGRS.2024.3379987.

[28] Q. Wei, D. Feng, and W. Jia, "UDR: An approximate unbiased difference-ratio edge detector for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6688–6705, Aug. 2021, doi: 10.1109/tgrs.2020.3031905.

[29] K. Mikolajczyk and C. Schmid, "indexing based on scale invariant interest points indexing based on scale invariant interest points," in *Proc. 8th IEEE Int. Conf. Comput. Vis.*, 2001, pp. 525–531, doi: 10.1109/ICCV.2001.937561.
[30] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *Int. J. Comput. Vis.*, vol. 60, no. 1, pp. 63–86, 2004, doi: 10.1023/B:VISI.0000027790.02288.f2.
[31] M. Pourfard et al., "KAZE-SAR: SAR image registration using KAZE detector and modified SURF descriptor for tackling speckle noise," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jun. 2022, Art. no. 5207612, doi: 10.1109/TGRS.2021.3084411.
[32] J. Li, Q. Hu, and M. Ai, "RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform," *IEEE Trans. Image Process.*, vol. 29, pp. 3296–3310, Dec. 2020, doi: 10.1109/TIP.2019.2959244.
[33] A. Nabatchian, E. Abdel-Raheem, and M. Ahmadi, "Illumination invariant feature extraction and mutual-information-based local matching for face recognition under illumination variation and occlusion," *Pattern Recognit.*, vol. 44, no. 10/11, pp. 2576–2587, 2011, doi: 10.1016/j.patcog.2011.03.012.
[34] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale-invariant feature matching for optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4516–4527, Nov. 2011, doi: 10.1109/TGRS.2011.2144607.
[35] H. Zhu, L. Jiao, W. Ma, F. Liu, and W. Zhao, "A novel neural network for remote sensing image matching," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2853–2865, Sep. 2019, doi: 10.1109/TNNLS.2018.2888757.
[36] H. Gonçalves, J. A. Gonçalves, and L. Corte-Real, "Measures for an objective evaluation of the geometric correction process quality," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 292–296, Apr. 2009, doi: 10.1109/LGRS.2008.2012441.

Zhonghua Hong (Member, IEEE) received the Ph.D. degree in GIS from Tongji University, Shanghai, China, in 2014.

He has been an Professor with the College of Information Technology, Shanghai Ocean University, Shanghai, since 2022. His research interests including satellite/aerial photogrammetry, high-speed videogrammetric, planetary mapping, 3-D emergency mapping, GNSS-R, deep learning, and processing of geospatial Big Data.

Yu Lu received the B.S. degree in information engineering from the Nanjing Institute of Technology, Nanjing, China, in 2021. He is currently working toward the M.S. degree in computer science and technology with Shanghai Ocean University, Shanghai, China.

His research interests include feature matching.

Yongsheng Geng received the B.S. degree in network engineering from the Tongda College of Nanjing University of Posts & Telecommunications, Nanjing, China, in 2019. He is currently working toward the M.S. degree in software engineering with Shanghai Ocean University, Shanghai, China.

His research interests include feature matching.

Xiaohua Tong (Senior Member, IEEE) received the Ph.D. degree in GIS from Tongji University, Shanghai, China, in 1999.

From 2001 to 2003, he was a Postdoctoral Researcher with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China. He was a Research Fellow with Hong Kong Polytechnic University, Hong Kong, in 2006, and a Visiting Scholar with the University of California, Santa Barbara, CA, USA, from 2008 to 2009. His research interests include photogrammetry and remote sensing, trust in spatial data, and image processing for high-resolution satellite images.

Shijie Liu received the Ph.D. degree in cartography and geographic information engineering from Tongji University, Shanghai, China, in 2012.

He is currently a Professor with the College of Surveying and Geo-informatics, Tongji University. His research interests include geometric exploitation of high-resolution remote sensing and its applications.

Ruyan Zhou received the Ph.D. degree in agricultural bio-environment and energy engineering from Henan Agricultural University, Zhengzhou, China, in 2007.

From 2007 to 2008, she worked with Zhongyuan University of Technology. Currently, she is working with Shanghai Ocean University, Shanghai, China.

Haiyan Pan received the Ph.D. degree in surveying and mapping from Tongji University, Shanghai, China, in 2020.

She has been a Lecturer with the College of Information Technology, Shanghai Ocean University, Shanghai. Her research interests include multitemporal remote sensing data analysis, change detection, and multispectral/hyperspectral image classification.

Yun Zhang received the Ph.D. degree in applied marine environmental studies from the Tokyo University of Maritime Science and Technology, Tokyo, Japan, in 2008.

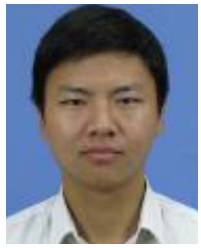
Since 2011, he has been a Professor with the College of Information and Technology, Shanghai Ocean University, Shanghai, China. His research interests include the study of navigation system reflection signal technique and its maritime application.

Yanling Han received the B.E. degree in mechanical design and manufacturing and the M.E. degree in mechanical automation from Sichuan University, Sichuan, China, in 1996 and 1999, respectively, and the Ph.D. degree in engineering and control theory from Shanghai University, Shanghai, China, in 2005.

She is a Professor and currently working with the Shanghai Ocean University, Shanghai. Her research interests include the study of ocean remote sensing, flexible system modeling, and deep learning.

Jing Wang received the Ph.D. degree in biomedical engineering from the Department of Biomedical Engineering, Shanghai Jiaotong University, Shanghai, China, in 2014.

She has been a Lecturer with the College of Information Technology, Shanghai Ocean University, since 2015. Her research interests include computer vision and medical image processing.



Shuhu Yang received the Ph.D. degree in physics from the School of Physics, Nanjing University, Nanjing, China, in 2012.

Since 2012, he has been the Lecturer with the College of Information Technology, Shanghai Ocean University, Shanghai, China. His research interests include the evolution of the Antarctic ice sheet, hyperspectral remote sensing, and the use of navigational satellite reflections.