A Robust Registration Method for Multi-view SAR Images based on Best Buddy Similarity

Yifan Zhang 1, Zhiwei Li 1, Wen Wang 2, Minzheng Mu 1, Bangwei Zuo 1

1 School of Geoscience and Info-physics, Central South University, 410083 Changsha, China
2 School of Electrical Engineering, Naval University of Engineering, 430000 Wuhan, China

Keywords: Image Registration, Synthetic Aperture Radar, Multi-View Image, Template Matching, Best Buddy Similarity.

Abstract

Due to the influence of imaging angle and terrain undulation, multi-view synthetic aperture radar (SAR) images are difficult to be directly registered by traditional methods. Although feature matching solves the issue of image rotation and maintains scale invariance, these methods often lead to non-uniformity of interest points and may not achieve subpixel accuracy. The traditional template matching method makes it difficult to generate correct matches for multi-view SAR oblique images. In this paper, a multi-view SAR image template matching method based on Best Buddy Similarity (BBS) is proposed to solve the traditional methods’ problem. Firstly, the initial correspondences between images are established according to the Range-Doppler model of SAR images. Secondly, a sliding window search is performed on the established correspondence, the BBS is calculated, and the subpixel locations of the peaks on the similarity map are estimated to achieve a fine match. In the calculation process of BBS, the SAR ROEWA operator is used to suppress the speckle noise of SAR images. The experiment demonstrated that SAR-BBS can accurately match SAR images with large rotation angle. The peak value on the search window is significant. The registration accuracy of SAR-BBS outperforms the other state-of-the-art methods.

1. Introduction

Synthetic aperture radar (SAR) image registration is a fundamental step required for many applications, and its quality directly affects the effectiveness of subsequent tasks. At present, there is a lot of research on SAR image registration algorithms, and they are applied to SAR interferometry (InSAR) deformation measurement (Zou et al., 2009), SAR offset tracking (Strozzi et al., 2002), and block adjustment tie-points matching (Huber et al., 2010). The raw SAR images are usually in the oblique direction, which is affected by the motion attitude of the satellite platform and the imaging angle (Xiang et al., 2022b). SAR images acquired on multi-view satellite platforms have various advantages, such as complementing each other with perspective shrinkage and overlapping problems that occur in undulating terrain areas (Jing et al., 2018). However, it also brings new challenges to SAR image registration algorithms.

SAR image registration methods can be roughly divided into three categories, pixel space-based matching methods, feature description-based matching methods, and deep learning methods (Sommervold et al., 2023). Among the traditional methods, the matching method based on pixel space is the most commonly used SAR image registration method, which accurately estimates the offset information between images by selecting the template, searching for the offset (which can be carried out in the time domain or frequency domain), and fitting the subpixel position, which is simple in principle and has high matching accuracy of homologous images. Common search criteria include normalized cross-correlation (NCC) (Pratt, 1978), sum of squared differences (SSD) (Mahmood and Khan, 2012), structure similarity index measure (SSIM) (Viola and Wells III, 1997), mutual information (MI) (Oliveira and Tavares, 2014), and other methods. However, the shortcomings of the pixel space-based matching method are also obvious, and the initial correspondences need to be given in advance, and the calculation time complexity is high (Sommervold et al., 2023). The feature-matching method needs to select interest points according to certain rules, use a set of descriptors, and find the correct matches through a certain similarity measure (Li et al., 2020). Most of the newly proposed feature matching methods are developed based on classical SIFT (Lowe, 2004), such as SAR-SIFT (Dellinger et al., 2015), SURF (Bay et al., 2008), PCA-SIFT (Ke and Sukthankar, 2004), ORB (Rublee et al., 2011), etc., which can achieve good matching effects for images with rotation, lighting, scale, etc., and images without prior information, but non-negligible mismatching and unevenly distributed interest points bring challenges to the method based on feature description matching. As an emerging SAR image registration method, deep learning has significant advantages in automatic operation and high efficiency (Sommervold et al., 2023). From CNN (Alzubaidi et al., 2021), U-Net (Siddique et al., 2021) to GAN (Strite and Morkoç, 1992), most methods extract features from the network and evaluate feature similarity, and the results are better than traditional feature matching accuracy. However, due to the training requirements of large samples, deep learning methods are difficult to ensure the robustness and timeliness of matching results.

Ghannadi et al. (Ghannadi et al., 2020) proposed to extract statistical, frequency, and polarized texture features before image registration, construct the second optimal texture image using the weighted average of extracted features, optimize the weight coefficient by using the cat swarm algorithm, and input the results into SURF for matching, which increases the number of real matches and improves uniformity. Pallotta et al. (Pallotta et al., 2022) proposed a constrained least squares method to eliminate outliers in multitemporal SAR registration. Xiang et al. (Xiang et al., 2022a) proposed a geometric perception method to mask the overlapping and shadowed areas of SAR images after geocoding to maintain matching accuracy. Huang et al. (Huang et al., 2022) proposed a registration algorithm for the segmentation concept to reduce the geometric distortion and grayscale inconsistency between dual-band SAR. Wang et al. (Wang et al., 2022) applied SAR-SIFT to the image-matching problem in the large-scale uncontrolled orthorectification method of GF-3 images in China, and parallelized the process of computation.
Although the SAR image registration method above has achieved good registration accuracy, the existing methods still face great challenges for multi-view SAR image registration. Geometric and radiometric differences caused by different-angle imaging are key issues in multi-angle image matching. Although the feature matching method can ensure the scale and rotation invariance to a certain extent, it is not possible to extract high-quality interest points in all regions, and it is difficult to meet the uniformity requirements of the control points of image registration. The template matching method focuses on the grayscale properties of the image itself. The control points are evenly distributed and the matching accuracy is high. However, the existing template matching method still has great limitations for multi-view SAR image registration.

In this paper, a robust method for multi-view SAR image matching is proposed. Firstly, for two multi-view oblique SAR images, the initial correspondences are generated using the provided satellite position information. Secondly, SAR-BBS based on the SAR-ROEWA operator is defined and used for the refined search process of correspondences. which can separate homogeneous structures with the same distribution in the mixed mode of speckle noise and image signal. Finally, fitting by polynomial model can achieve accurate registration of two-scene multi-view SAR images.

2. Methodology

2.1 Overall process

The registration method proposed in this paper consists of two stages. The overall process is shown in Figure 1. Firstly, the initial correspondences on reference and sense image are generated by regular grid division, which is the same as most template matching methods. The interest point is connected to the counterpart on the sense image by the RD model which uses external DEM.

![Figure 1. Processing of proposed method which consists of two stages.](image)

Secondly, the template matching based on SAR-BBS for each correspondence is proposed. A SAR-Harris map based on the SAR-ROEWA operator is generated. Then the template of the point is extracted and is used to calculate the number of best buddy point (BBP) for correspondence. With the template swiping on Sense Image, every BBP on the search space is calculated. BBS is the collection of BBP in the search space. Finally, we searched the correlation peak on the BBS to get the most appropriate offset.

2.2 Initial correspondences generation

For the two SAR images obtained in different orbits, the interest points are usually matched through the image orbit parameter file provided by the satellite transmitter. The original SAR image is in a radar coordinate system, also known as a radar-Doppler coordinate system (Luo et al., 2022). The position of ground point P on the radar-Doppler coordinate system is expressed as

\[ P = [X(t), Y(t), Z(t)] \]

Based on the SAR imaging equation, the following relationship is obtained.

\[
\begin{align*}
R &= |P(X,Y,Z) - S| \\
f_{\text{dop}} &= \frac{2V}{\lambda} \cdot |P(X,Y,Z) - S|
\end{align*}
\]

where \( S \) and \( V \) are the position and velocity vectors of the satellite in the space Cartesian coordinate system at \( t_0 \) moment, \( \lambda \) is the working wavelength of the SAR sensor, \( f_{\text{dop}} \) is the Doppler frequency of the SAR sensor, and \( a, b, \) and \( h \) represents the major and semi-major axes, short semi-axes and earth height of the reference ellipsoid, respectively.

In secondary image imaging, whenever the time \( t_0 \) changed by \( \Delta t \) relative to the imaging moment, it is iteratively calculated in the following equation.

\[
m_{\text{in}} = \sqrt{(X - X(t))^2 + (Y - Y(t))^2 + (Z - Z(t))^2}
\]

In Eq. (2), \( X(t), Y(t), \) and \( Z(t) \) are the positions of the satellite in the Cartesian coordinate system of space at time \( t \), respectively. Using the position-time relationship in the satellite orbit data, \( X(t), Y(t), \) and \( Z(t) \) can be solved by polynomial fitting parameters. The fitted model is

\[
\begin{align*}
X(t) &= a_0 + a_1t + a_2t^2 + a_3t^3 \\
Y(t) &= b_0 + b_1t + b_2t^2 + b_3t^3 \\
Z(t) &= c_0 + c_1t + c_2t^2 + c_3t^3
\end{align*}
\]

Based on Eq. (1), (2), the position of \( P \) on the secondary image coordinate system \( P_2(r_j, c_j) \) can be calculated by

\[
\begin{align*}
X(t) &= a_0 + a_1t + a_2t^2 + a_3t^3 \\
Y(t) &= b_0 + b_1t + b_2t^2 + b_3t^3 \\
Z(t) &= c_0 + c_1t + c_2t^2 + c_3t^3
\end{align*}
\]

where \( S_j \) is the current position of the satellite calculated in Eq. (3), \( R_2 \) is the smallest oblique value on the SAR Range-Doppler coordinate system, and \( \Delta R \) is the progressive step of the oblique range. Based on the Range-Doppler model, a rough offset from the reference image point to the secondary image can be obtained.

2.3 Template matching process

After estimating the coarse offset, some correspondences can be extracted from the imagery. On this basis, the template matching process is carried out to refine the correspondence of homogeneous points. The traditional template matching method makes it difficult to achieve high accuracy for multi-view SAR images at this time. Therefore, we improve the similarity of BBS (Oron et al., 2018) and make it suitable for SAR image matching.
according to the statistical characteristics of SAR noise. The BBS method calculates the number of BBP in two points sets. When the BBS application matches the sliding window template, the template content is considered to have the greatest similarity when the number of BBP in the two points sets is the peak in the entire search area. BBS is defined as follows.

$$BBS(P, Q) = \frac{1}{\min(M, N)} \sum_{i=1}^{N} \sum_{j=1}^{M} bb(p_i, q_j, P, Q). \quad (5)$$

where $P$ and $Q$ are part of the points set of the template image and the search window, respectively, expressed as $P = \{ p_i \}_{i=1}^{N}$ and $Q = \{ q_j \}_{j=1}^{M}$, $M$ and $N$ represent the number of elements of the points set $P$ and $Q$, respectively. $bb$ is a pair of point $\{ p_i \in P, q_j \in Q \}$ satisfies the expression of the best buddy

$$bb(p_i, q_j, P, Q) = \begin{cases} 1 & NN(p_i, Q) = q_j \wedge NN(q_j, P) = p_i, \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where $NN(p_i, Q)$ means that $p_i$ in points set $P$ is the nearest neighbour of the $q_j$ in points set $Q$, which satisfies

$$NN(p_i, Q) = \text{argmin}_q d(p_i, q). \quad (7)$$

$d(p_i, q)$ in Eq. (7) is a distance function similar to Euclidean distance, cos distance, etc. The distance function used by Dekel et al. is

$$d(p_i, q_j) = \|p_i^{(A)} - q_j^{(A)}\|_2^2 + \lambda \|p_i^{(L)} - q_j^{(L)}\|_2^2. \quad (8)$$

where $A$ represents the pixel value of the point, $L$ represents the location of the point, and for the constant $\lambda$, we use the suggested empirical value that $\lambda = 2$. Since BBS is sensitive to noise, in SAR-BBS, we processed the contents of the points set $P$ in Eq. (7). Since the noise model of SAR images is multiplicative, the ROEWA operator is used to filter the input points set in the horizontal and vertical directions (Xiang et al., 2018), respectively. ROEWA is expressed as

$$\text{ROEWA}_h = \frac{\sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} I(x+mn)e^{-|\ln|x+y|\alpha}}{\sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} I(x+mn)}$$

$$\text{ROEWA}_v = \frac{\sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} I(x+mn)e^{-|\ln|x+y|\alpha}}{\sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} I(x+mn)} \quad (9)$$

where $M, N$ represent the number of points in the two-dimensional points set, $\alpha$ is the spatial scale parameter, and $I(x, y)$ represents the pixel value of the current point at the coordinates on the image. Based on this, the gradients in the horizontal and vertical directions are expressed as

$$G_h = \log(\text{ROEWA}_h, \alpha)$$

$$G_v = \log(\text{ROEWA}_v, \alpha) \quad (10)$$

After achieving the gradient extraction and noise suppression of the points sets $P$ and $Q$, the down-sampling operation is performed spatially to reduce the computational overhead in the process of calculating the number of BBP by window sliding.

3. Results

3.1 Experimental datasets and environments

In this paper, two TerraSAR images obtained at different angles and different orbits are used to verify the matching accuracy of the proposed SAR-BBS algorithm. The experiment is divided into two parts. The first is the template similarity experiment and analysis, which is compared with similarity measures such as NCC and MI. The second is the matching accuracy analysis, which is compared with two state-of-the-art multi-view image registration algorithms: SAR-SIFT and RIFT. The relationship between the coverage area and the overlap of the real SAR image for the matching experiment is shown in Figure 2. Table 1 lists the parameter information of the images.

![Figure 2. Experimental area.](image)

<table>
<thead>
<tr>
<th>Item</th>
<th>Reference Image</th>
<th>Sense Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor</td>
<td>TerraSAR-1</td>
<td>TerraSAR-1</td>
</tr>
<tr>
<td>Date</td>
<td>2013–3–20</td>
<td>2012-10-13</td>
</tr>
<tr>
<td>Heading</td>
<td>189.83°</td>
<td>350.15°</td>
</tr>
<tr>
<td>Incident angle</td>
<td>39.32°</td>
<td>39.23°</td>
</tr>
<tr>
<td>Multilook ratio</td>
<td>Azimuth:5</td>
<td>Azimuth:5</td>
</tr>
<tr>
<td></td>
<td>range:5</td>
<td>range:5</td>
</tr>
<tr>
<td>Image size</td>
<td>6111×3050</td>
<td>6111×3050</td>
</tr>
<tr>
<td>Centre position</td>
<td>Geo lat,lon</td>
<td>Geo lat,lon</td>
</tr>
</tbody>
</table>

Table 1. Experiment image information

3.2 Evaluation Criteria

In the experiments, correct matching ratio (CMR) and root mean square error (RMSE) were used to quantitatively evaluate the results of the registration method. Here, CMR is defined as
\[ CMR = \frac{CM}{TM} \]  

where \( CM \) represents the correct number of matches and \( TM \) represents the total number of matches. The correct matches are calculated by the fast sample consensus (FSC) algorithm (Wu et al., 2015).

### 3.3 Comparison of template matching similarity

Firstly, we apply different template-matching similarity measures to some regions in the real SAR images to test the effectiveness of the SAR-BBS template matching proposed in this paper for SAR images with different viewing angles. The most widely used NCC and MI similarity measures were selected for comparative testing, and the results are shown in Figure 3.

According to the results shown in Figure 3, the selected template region has a large angular deformation and plane displacement, which brings great challenges to the similarity measure of template matching. For the template on the reference patch, a sliding window search is performed on the sensing patch, and the similarity information calculated on the search window forms a similarity map. The SAR-NCC map peaked on both test patches, which had the largest number of BBP in points sets and was considered the best match within the search area. NCC map performed poorly in search results on the first patch, did not show a unique peak, and was mixed with multiple low NCC value noise peaks, which means that the NCC method failed in this region. The NCC method has two peaks on the second test patch, one is a relatively stable true peak, and the other peak is at the edge of the window, which may be due to the pseudo peak caused by the matching boundary. For the first true peak, it is easy to see how much distortion the image is observed, and it does not reflect the true offset position of the image. For the MI method, both test patches fail and the similarity map does not show an exact matching peak point. NCC and MI methods rely on the value of image pixels, and it is difficult to match the correct correspondence for images with rotation and local structural distortion. The SAR-BBS method uses the ROEWA operator to suppress SAR images with speckle noise, and adopts a more robust BBP expression for image structure information, which can successfully cope with image matching with large rotation angle and viewing angle distortion.

### 3.4 Comparison with state-of-the-art

The qualitative results compared with SAR-SIFT and RIFT are shown in Figure 4-6. In Figure 4, the SAR-SIFT method is too focused on the extraction of interest points, which causes the local deformation information to be enlarged to the overall image area when the transformation model is fitted, resulting in unnecessary registration deformation. Figure 5 shows that the RIFT method is still challenging in matching SAR images obtained for different viewing angles, and the coarse matching results without error point checking have more errors. As shown in Figure 6, the SAR-BBS method has a uniformly distributed matching number because it regularly selects grid points distributed under the SAR image coordinate system. The matching results of SAR-SIFT and RIFT are coarse differences rejected by the FSC method. The transformation model is set as affine. The quantitative experimental results are shown in Table 2.
distribution through the distance relationship between the points sets, and increases the robustness by sliding the window. However, the BBS method is very complicated in calculating the distance of the points set, which also brings great time consumption.

### 4. Conclusion

In this paper, a robust registration method is proposed for SAR images acquired under multi-viewing conditions, and the accuracy and effectiveness of the proposed method are highlighted by comparing it with state-of-the-art and commonly used methods. The contributions made in this article are as follows. SAR-BBS was proposed and applied to the registration of multi-view SAR oblique images. The ROEWA operator was used to suppress the interference of SAR spot noise on the matching template. Image templates are converted to points sets, and the effect of attenuating rotation on image matching. BBP is calculated to count the number of pixels with the same distribution form. The template-matched sliding window is used to increase the robustness of BBP to noise. By comparing with NCC and MI methods, the effectiveness of BBS on rotating templates is verified. The comparison with SAR-SIFT and RIFT shows that SAR-BBS has higher matching accuracy and a reliable number of points for multi-view SAR images.

### Acknowledgements

This paper is funded by the Hunan Provincial Innovation Foundation for Postgraduate under Grant CX20230203, and in part by the Fundamental Research Funds for the Central Universities of Central South University under Grant 2023zzts0119.

### References


Based on the experimental results in Table 2, it is obvious that SAR-BBS outperforms other methods at CMR and RMSE. As a template matching method, SAR-BBS can initially control the interest point according to the correspondence established by the Range-Doppler model, and then use the sliding window to search for the corresponding peak value, so as to refine the matches. BBS processes the template image as points set, reduces the dependence on the pixel position information in the image, extracts a subset with a statistically higher probability of the same

<table>
<thead>
<tr>
<th>Item</th>
<th>SAR-SIFT</th>
<th>RIFT</th>
<th>SAR-BBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial matches</td>
<td>145</td>
<td>214</td>
<td>106</td>
</tr>
<tr>
<td>Correct matches</td>
<td>40</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>CMR (%)</td>
<td>27.5</td>
<td>2.80</td>
<td>28.31</td>
</tr>
<tr>
<td>RMSE (pixel)</td>
<td>1.635</td>
<td>1.829</td>
<td>1.046</td>
</tr>
<tr>
<td>Time (s)</td>
<td>67.173</td>
<td>33.674</td>
<td>104.218</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>SAR-SIFT</th>
<th>RIFT</th>
<th>SAR-BBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial matches</td>
<td>145</td>
<td>214</td>
<td>106</td>
</tr>
<tr>
<td>Correct matches</td>
<td>40</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>CMR (%)</td>
<td>27.5</td>
<td>2.80</td>
<td>28.31</td>
</tr>
<tr>
<td>RMSE (pixel)</td>
<td>1.635</td>
<td>1.829</td>
<td>1.046</td>
</tr>
<tr>
<td>Time (s)</td>
<td>67.173</td>
<td>33.674</td>
<td>104.218</td>
</tr>
</tbody>
</table>

Table 2 Comparison of the relative accuracy of registration


