The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-1/W2-2023
ISPRS Geospatial Week 2023, 2–7 September 2023, Cairo, Egypt

MULTI-ORIENTATION EDGE-BASED SATELLITE IMAGE MATCHING METHOD FOR OPTICAL AND SAR IMAGES

Yi Lin [1 2], Lang Li [1], Lujia Wei [1], Liming Liu [3], Jie Yu [1 *]

[1] College of Surveying and Geo-Informatics, Tongji University, 200092 Shanghai, China – (linyi, lilang, 2131949, 2011_jieyu)@tongji.edu.cn
[2] Research Center of Remote Sensing Technology and Application, Tongji University, 200092 Shanghai, China
[3] Shanghai Ubiquitous Navigation Technology Co., Ltd., 201702 Shanghai, China – llm@ubinavi.com.cn

KEY WORDS: Satellite Image Matching, Optical-SAR Image Pairs, Edge Detector, Multiple Orientation Edge Feature Correlation, Nonlinear Radiance Distortion.

ABSTRACT:

Multi-sensor image matching is a key technology to fully exploit complementary information on images from multiple sources. In particular, the nonlinear radiance distortion between optical and SAR images makes image matching very difficult. Focusing on satellite optical and SAR image matching, this paper proposed an oriented edge-based template image matching, namely Multiple Orientation Edge Feature Correlation (MOEFC). Firstly, the edge feature was extracted by the Sobel operator to construct oriented gradient channels on optical and SAR images. Then spatial filtering was performed separately by a two-dimensional and a one-dimensional kernel. Finally, the similarity was determined by stacking the normalized cross correlation (NCC) of each oriented gradient channel. To validate the new method, it was compared with other four edge-based template image matching methods detected by Sobel, Canny, Laplacian and phase congruency-based methods on satellite optical-SAR image pairs, which were tested in rural and urban areas. And the results show that the proposed MOEFC method can obtain the largest number of correct matching (NCM) with the least root mean square error (RMSE) in both areas. In the rural area, the NCM of the MOEFC method is higher than the other four methods, and the improvement ranges from 54.1% to 74.3%. And the RMSE of the MOEFC method is 0.645 pixel. In the urban area, the NCM of the MOEFC method can also be improved by 3.6% to 46.4%. And the RMSE of the MOEFC method equals 0.489 pixel.

1. INTRODUCTION

In recent years, with the increasing demand for Earth Observation and technological advances, the ability of various types of sensors to acquire data has continued to improve, and the spatial resolution, temporal resolution, spectral resolution and the coverage of remote sensing images have continued to increase (Sui et al., 2022). Accurate image matching is a key step for applications such as information registration, fusion, 3D reconstruction, object recognition and change detection for heterogeneous remote sensing images (Dawn et al., 2010). Although image matching is a well-known problem, it is still a challenging task, especially for optical and SAR images. The nonlinear radiance distortion between two image types makes image matching very difficult. In addition, the inherent speckle noise degrades the image quality, which increases the difficulty of image matching (Yao et al., 2021). Therefore, a robust method that can adapt to the nonlinear radiance distortion and are insensitive to the SAR image noise is an urgent and necessary need.

Image matching is widely divided into 3 strategies: template-based matching, feature-based matching and deep learning-based matching. Feature-based matching extracts salient features and then matches these features according to their similarities (Xu et al., 2016). Recent studies have made great progress in which using local-invariant features (Xiang et al., 2018) and phase-based features (Fan et al., 2018; Li et al., 2020). But the need for highly repeatable features between image pairs still affects the performance of feature-based methods (Ye et al., 2017). Deep learning-based matching automatically extracts high-level features that are less sensitive to nonlinear radiance distortions and significant geometric distortions (Lan et al., 2021; Ye et al., 2022). However, the availability of training samples limits the generalization capacity in different application scenarios (Zhang et al., 2023). Template-based matching is also known as area-based image matching, which needs to determine the similarity between the template from the referenced image and the matched image in the spatial or frequency domain (Zhu et al., 2022). However, traditional template image matching usually relied on the intensity of images which has been found to be hardly applied to the optical and SAR image matching research due to the nonlinear radiance distortion (Ye et al., 2017). To overcome the intensity dependence issue, structural similarity has been used to depict the contours of multi-sensor images (Wang et al., 2022).

Since the structural information is usually associated with the edge features, it is natural to use the edge feature for template image matching. Edge-based template image matching aims to find the location where the gradient modulus of the template is closest to that in the search region (Carsten et al., 2008). There are many edge detection methods (Jing et al., 2022) that can be used in the edge-based template image matching task. The main problem is how to select an optimal edge feature for optical-SAR image matching that is insensitive to the nonlinear radiance distortion.

This paper proposed an oriented edge-based template image matching method, namely Multiple Orientation Edge Feature Correlation (MOEFC), which can resist the nonlinear radiance distortion. The core of MOEFC is to determine the similarity of the optical and SAR image pairs at different orientations instead

* Corresponding author

of using the edge intensity only. The n-oriented gradient channels on optical and SAR image pairs were first constructed followed by two times spatial filtering operations. Then, the normalized cross correlation (NCC) was used to determine the similarity of each orientation by sliding the template over the search region. The final similarity map was formed by stacking the orientation similarity map.

2. METHODOLOGY

In this section, the technological route is presented and the principle of the proposed MOEFC is explained in detail. In addition, two edge-based template image matching methods used in this study are illustrated as well.

2.1 Technical route of image matching for optical and SAR images

The technical route (Figure 1) includes five steps, and the detailed description is illustrated as follows:

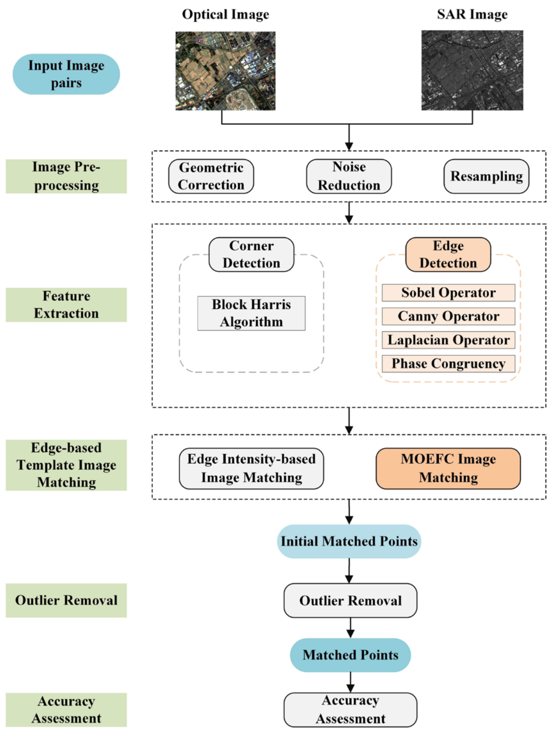


Figure 1. Technical Flowchart.

Step 1: Image pre-processing. The main purpose of this step is to reduce the geometric distortion of image pairs, but there are still nonlinear radiance distortions between different types of sensors. The detailed process consists of geometric correction, denoising and resampling. The Frost filter was used for noise reduction on SAR images to improve the image quality.

Step 2: Feature extraction. The corner detection was performed by the block Harris algorithm (Ye et al., 2017) on the optical image. And the edge detection process was conducted by four widely used edge detectors (Sobel operator, Canny operator, Laplacian operator and phase congruency method) on both optical and SAR images.

Step 3: Edge-based template image matching. There are two types of edge-based template image matching. The main difference between the edge intensity-based image matching method and the MOEFC image matching method is whether to calculate the contribution of intensity variations from different orientations. Four kinds of edge features extracted in step 2 are used in the edge intensity image matching and the edge feature extracted by the Sobel operator is used in the MOEFC method. And the description of the edge-based image matching method is given in section 2.3.

Step 4: Outlier removal. The random sample consensus (RANSAC) algorithm (Fischler and Bolles, 1981) was used to remove the initial matched points whose error was greater than 1 pixel.

Step 5: Accuracy assessment. All five edge-based methods were evaluated in both the qualitative and the quantitative way. The root mean square error (RMSE) and the number of correct matching (NCM) were selected as the indices for the quantitative assessment.

2.2 The construction of MOEFC

The MOEFC method is inspired by the Channel Features of Oriented Gradient (CFOG) design (Ye et al., 2019). It extracted horizontal and vertical edges using two one-dimensional Sobel operators, which are written as I_x, I_y. Then oriented gradient channels of number n were constructed. Each of the oriented gradient channels was determined by projecting the gradient along the *i*-th orientation, corresponding to the angle $\alpha = (i-1)\cdot 2\pi / \mathrm{n}, i = 1, 2, \cdots, n$. To deal with the gradient reversal, the *i*-th oriented gradient channel g_i is written as:

$$g_i = \left|\cos\alpha \cdot I_x + \sin\alpha \cdot I_y\right| \tag{1}$$

where $|\cdot|$ stands for obtaining the absolute value.

The n-oriented gradient channels were then smoothed in the spatial domain to reduce the gradient direction distortions. The smoothing operation consisted of a two-dimensional filtering on each oriented gradient channel and a one-dimensional filtering across the channel, which can be expressed as:

$$\begin{aligned} g_{i,\sigma} &= g_i * G(x, y) \\ g_{i,\sigma} &= 0.25 g_{i+1,\sigma} + 0.5 g_{i,\sigma} + 0.25 g_{i-1,\sigma} \end{aligned} \tag{2}$$

Where $*$ stands for the convolution operator in the spatial domain, $G(x, y) = \frac{1}{2\pi\sigma^2} \mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}$ is the two-dimensional Gaussian kernel function.

In this study, the NCC was used to determine the similarity of each oriented gradient channel. Then the final similarity map was created by stacking the n similarity maps. And this process can be displayed in Figure 2.

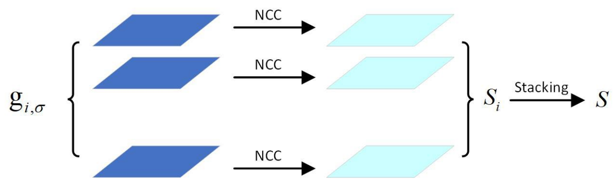


Figure 2. Schematic diagram of MOEFC.

2.3 Edge-based template image matching method

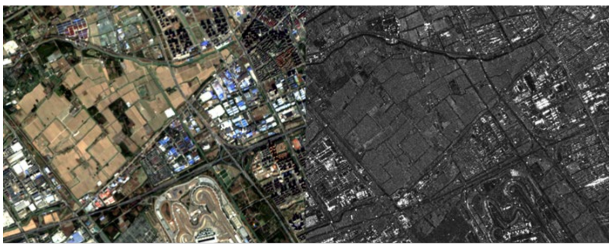
Edge-based template image matching of optical-SAR image pairs consisted of 3 parts: template extraction, search strategy and similarity calculation. First, templates were established on the edge feature of the optical image. Each template was a local window centered at a corner. Then a search region with a larger size than the template on the edge of the SAR image was determined. After that, the template was moved over the search region to find the overlapping region with the greatest similarity. For the similarity measurement, the NCC was directly used in the edge intensity-based image matching method. And the sum of NCC along all the oriented gradient channels was used in the proposed MOEFC method. Finally, the central pixels of the overlapping region with the highest similarity value were chosen as the initial matched points.

3. EXPERIMENT AND ANALYSIS

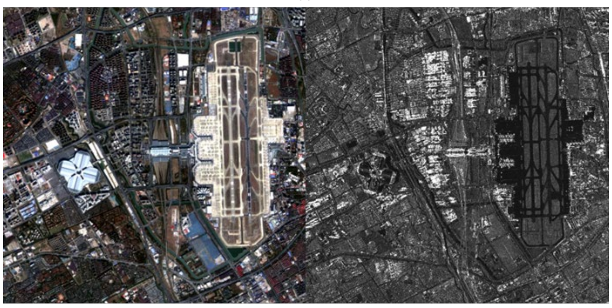
In this section, edge features extracted by the Sobel operator, Canny operator, Laplacian operator and phase congruency method were used in the edge intensity-based image matching, for comparison with the proposed MOEFC method in the rural and urban area. All the image matching methods were executed in Visual Studio 2017 using C++ programming.

3.1 Data source

Optical-SAR image pairs of two test areas with different image sizes, different optical data sources, and different land use and cover types were used in the experiments. Optical images were multispectral imagery acquired by the ZY1E and Gaofen-1 (GF-1) satellites. And the SAR images were both from TerraSAR-X in stripmap mode. To avoid the influence of seasonal changes on the image matching results, all the images were acquired in the winter of 2020. The detailed description is shown in Table 1. Image pair 1 was resampled to 10 m, and image pair 2 was resampled to 8 m (Figure 3).



(a)



(b)

Figure 3. Satellite optical-SAR image pairs for image matching. (a) Rural area: ZY1E (Left), TerraSAR-X (Right) (b) Urban area: GF-1 (Left), TerraSAR-X (Right).

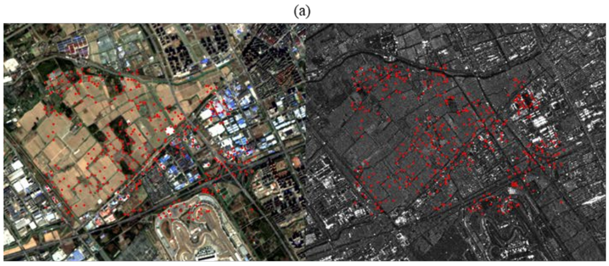
No.	Source	Size (H×W)	Characteristics
1	ZY1E TerraSAR-X	449×571	Rural area, mainly occupied by farmland
2	GF-1 TerraSAR-X	817×872	Urban area, mainly occupied by building land

Table 1. Details of the satellite image pairs for experiment.

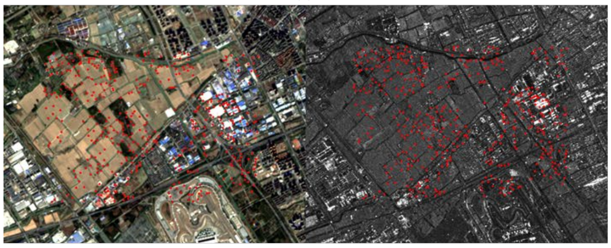
3.2 Results and analysis of the rural area

All five image matching methods were applied to image pair 1 to compare their effectiveness in the rural area from a qualitative and quantitative point of view.

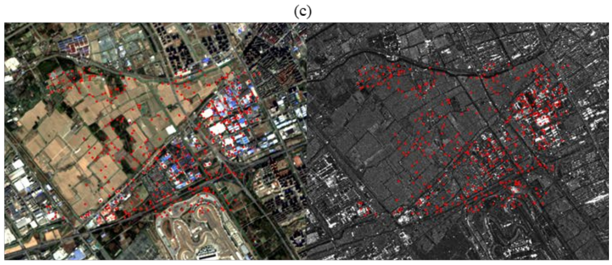
(a)



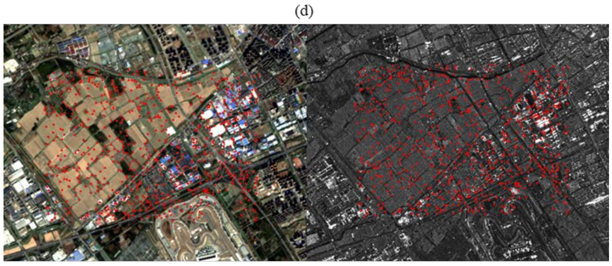
(b)



(c)



(d)



(e)

Figure 4. Different edge-based image matching results in rural area. (a) Sobel-based (b) Canny-based (c) Laplacian-based (d) Phase congruency-based (e) MOEFC.

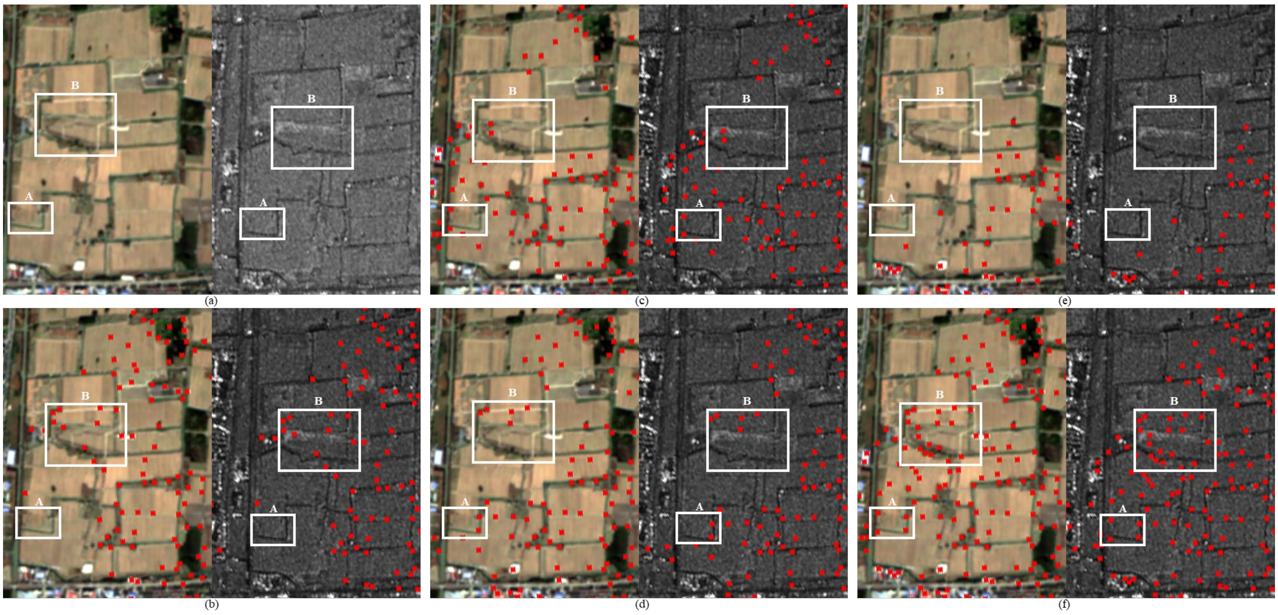


Figure 5. Zoom in results of different image matching methods in the farmland dominated area of image pair 1. (a) Subset of image pair 1 (b) Sobel-based (c) Canny-based (d) Laplacian-based (e) Phase congruency-based (f) MOEFC.

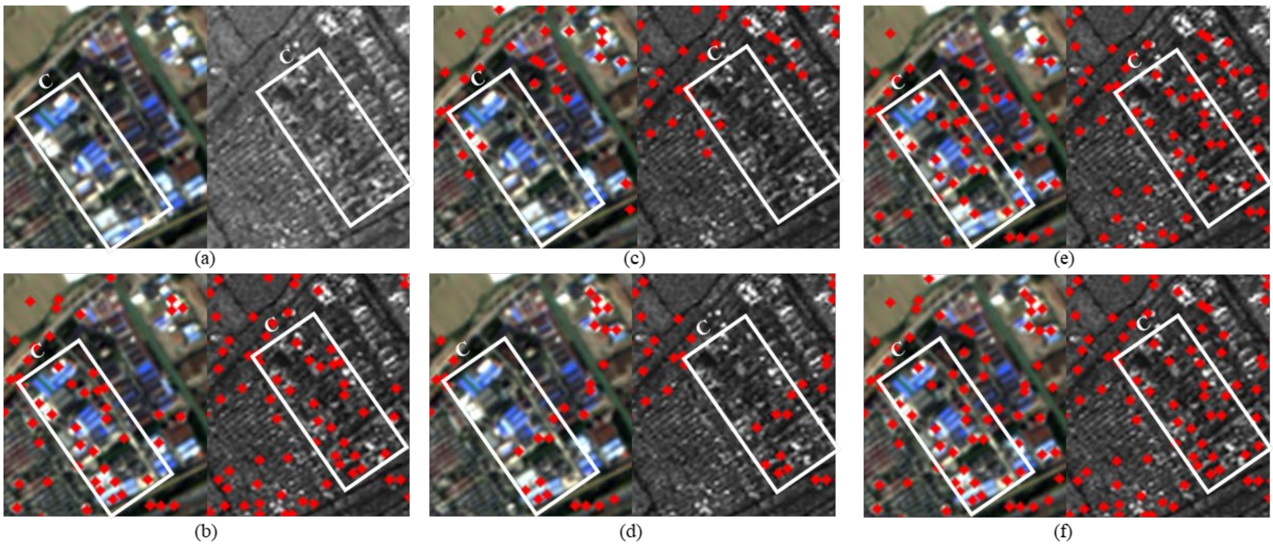


Figure 6. Zoom in results of different image matching methods in the building land dominated area of image pair 1. (a) Subset of image pair 1 (b) Sobel-based (c) Canny-based (d) Laplacian-based (e) Phase congruency-based (f) MOEFC.

Figure 4 shows the result of image matching. From the qualitative view, the matched points have different spatial distribution characteristics. In general, the proposed MOEFC method is able to find the most densely and evenly distributed matched points than the other four edge-based image matching methods. But the edge-based template image matching detected by the Sobel operator (Sobel-based), Canny operator (Canny-based), Laplacian operator (Laplacian-based) and phase congruency-based methods show different features in different subregions of Figure 4. Specifically, subregions A and B in Figure 5 indicate that the Sobel-based, Canny-based and Laplacian-based methods can obtain more matched points than the phase congruency-based method in the areas dominated by the farmland. However, the MOEFC method can obtain more matched points, especially in the left part of subregion A and along the edges of subregion B. In addition, subregion C in Figure 6 shows that the Sobel-based, Laplacian-based and phase congruency-based methods can obtain more matched points than the Canny-based method in the area dominated by the building land. And the MOEFC method also displays a comparable image matching ability to that of Sobel-based and phase congruency-based methods.

The NCM and RMSE of each method are plotted in Figure 7. It shows that the MOEFC method finds the largest NCM with imaging matching errors less than 1 pixel. It finds a total of 706 points. Compared with Sobel-based, Canny-based, Laplacian-based and phase congruency-based methods, the ratio of NCM increases by 55.5%, 74.3%, 54.1% and 55.2%, respectively. The RMSE of the MOEFC is 0.645 pixel while the RMSE of Sobel-based, Canny-based, Laplacian-based and phase congruency-based methods are 0.646 pixel, 0.702 pixel, 0.820 pixel and 0.810 pixel. Although the RMSE of Sobel-based method and MOEFC method are approximate, the NCM of the

https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-1417-2023 | © Author(s) 2023. CC BY 4.0 License.

MOEFC method validates that using the sum of oriented gradient similarity is more suitable for optical-SAR image matching in the rural area.

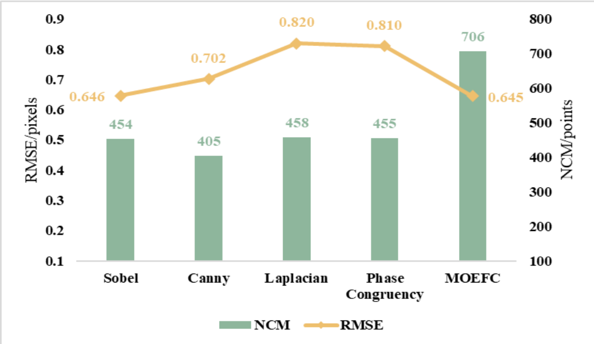


Figure 7. Image matching results in the rural area.

3.3 Results and analysis of the urban area

Figure 8 shows the result of image matching in the urban area using image pair 2. In terms of spatial distribution, MOEFC and Canny-based methods acquire more uniformly distributed matched points. However, the proposed MOEFC method is robust in different subregions. Subregion D in Figure 9 displays that the Sobel-based and Laplacian-based methods lack matched points in the upper left part and bottom part separately. And the phase congruency-based method can only obtain sparse matched points in subregion D. However, the MOEFC method and the Canny-based method can obtain more matched points in the same subregion. In Figure 10, the Laplacian-based method does not find the matched points at the bottom of subregion E. The Sobel-based, Canny-based and Laplacian-based methods acquire fewer matched points along the edges than the MOEFC method and the phase congruency-based method.

The NCM and RMSE of each method are recorded in Figure 11. It also confirms that the MOEFC method finds the largest NCM with imaging matching errors less than 1 pixel in the urban area. Specifically, the MOEFC method recognizes 2851 points which ranks first among the five methods. And the NCM of the MOEFC method is 17.2%, 3.6%, 46.4% and 27.7% higher than that of Sobel-based, Canny-based, Laplacian-based and phase congruency-based methods. Besides, the RMSE of the MOEFC is 0.489 pixel. But the RMSE of Sobel, Canny, Laplacian and phase congruency-based method equals 0.504 pixel, 0.536 pixel, 0.586 pixel and 0.543 pixel.

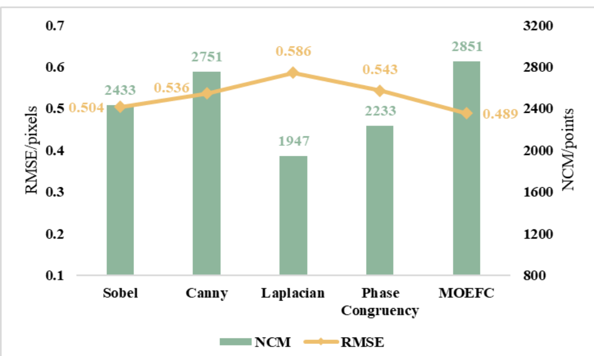


Figure 11. Image matching results in the urban area.

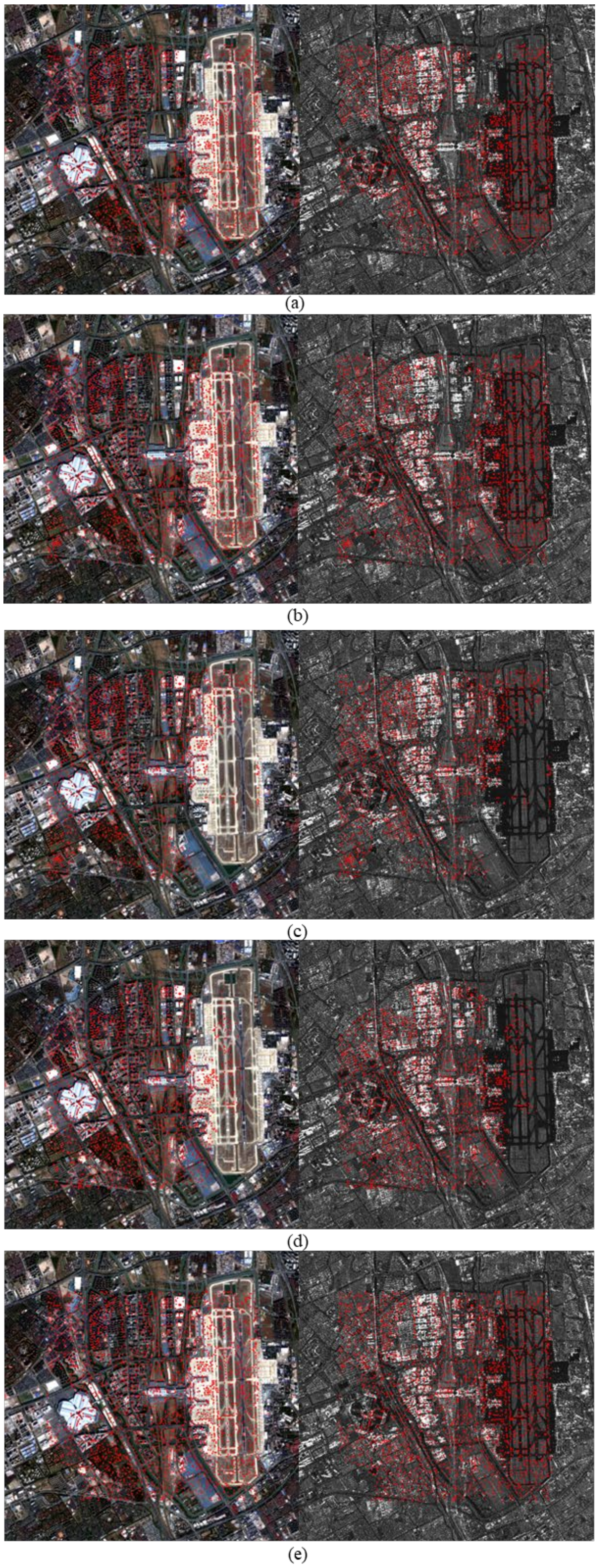


Figure 8. Different edge-based image matching results in urban area. (a) Sobel-based (b) Canny-based (c) Laplacian-based (d) Phase congruency-based (e) MOEFC.

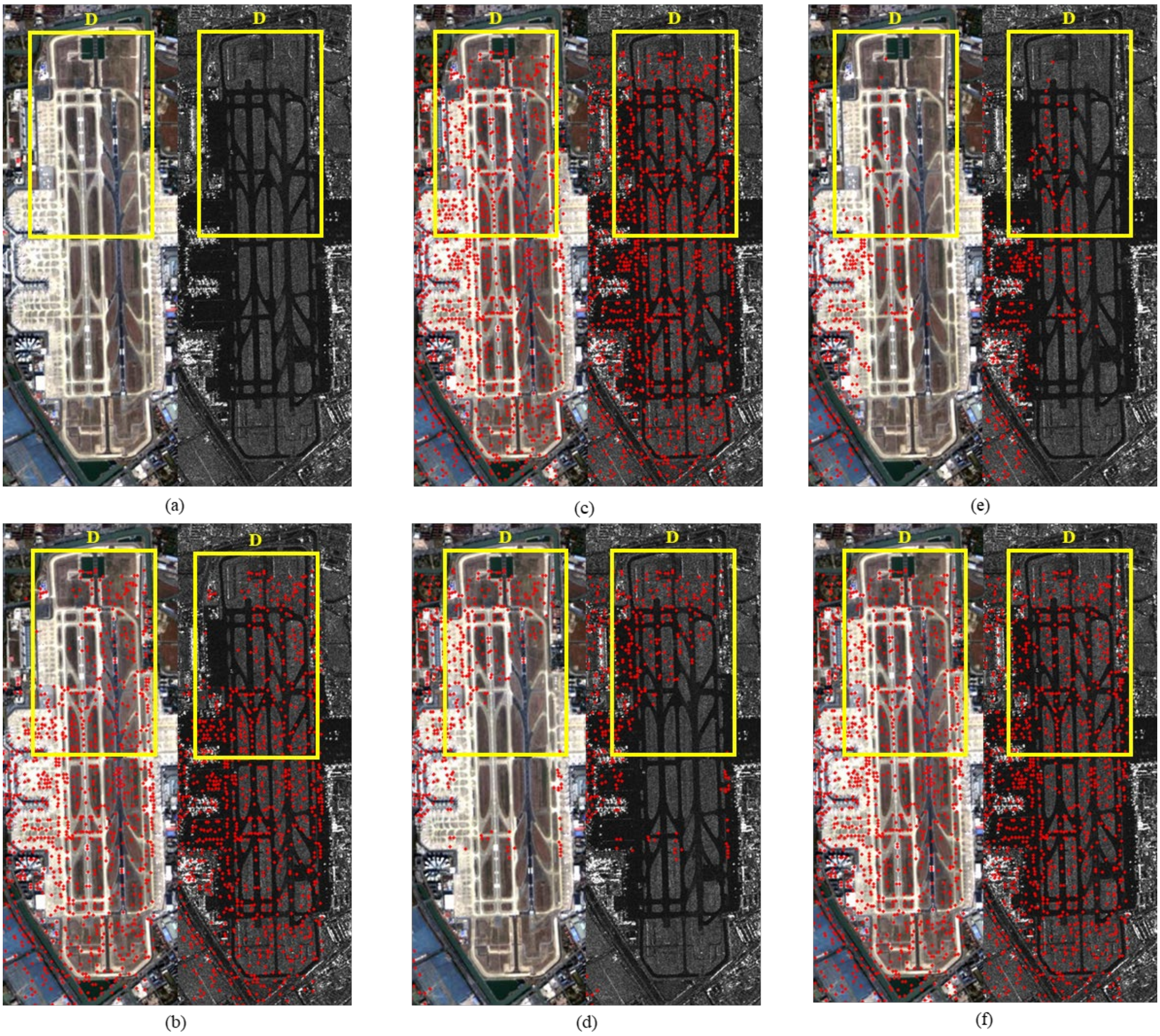


Figure 9. Zoom in results of different image matching methods on a large building dominated area of image pair 2. (a) Subset of image pair 2 (b) Sobel-based (c) Canny-based (d) Laplacian-based (e) Phase congruency-based (f) MOEFC.

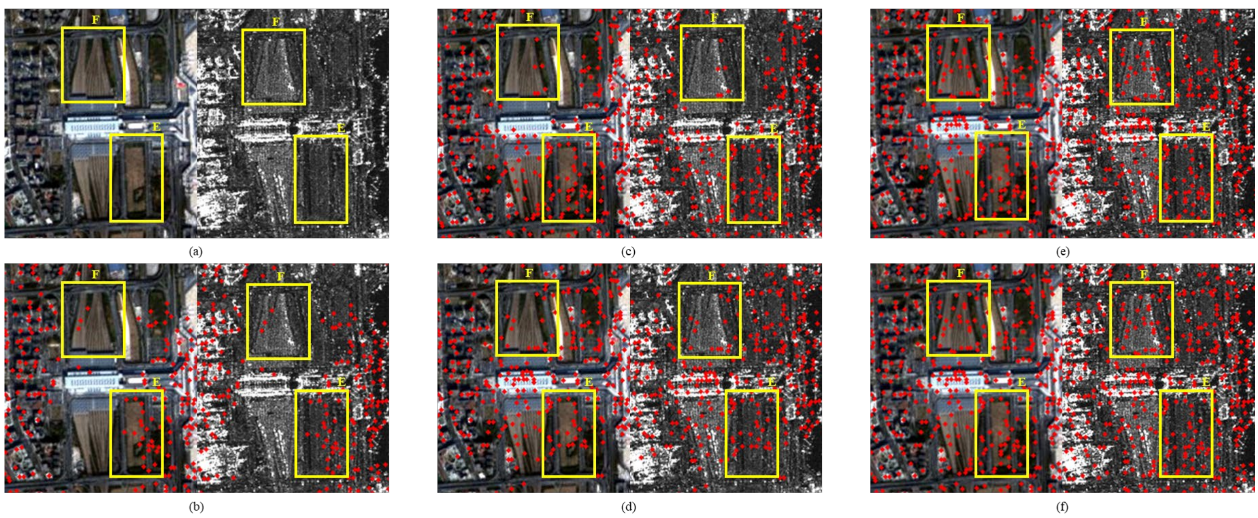


Figure 10. Zoom in results of different image matching methods in a building land area of image pair 2. (a) Subset of image pair 2 (b) Sobel-based (c) Canny-based (d) Laplacian-based (e) Phase congruency-based (f) MOEFC.

4. CONCLUSIONS

In this paper, a multi-orientation edge-based template matching method was proposed for satellite optical and SAR image matching. The NCC of each oriented gradient channel was stacked to determine the similarity of the optical-SAR image pair in the MOEFC method. Four edge intensity-based template image matching methods were utilized for comparison with the MOEFC method. Then, the image matching results were analyzed qualitatively and quantitatively, which helps to validate the effectiveness of the proposed method from different perspectives. According to the experimental results, the matched points obtained by the MOEFC are denser and more uniformly distributed in different subregions of rural and urban areas, especially in the farmland-dominated subregions of the rural area. But the spatial distribution characteristics of the other four edge-based template image matching methods may vary in different subregions. In terms of the quantitative assessment, the NCM of the MOEFC method increased by more than 54.1% in the rural area and by more than 3.6% in the urban area. In addition, the RMSE using the MOEFC method is 0.645 pixel and 0.489 pixel in the rural and urban areas, respectively. And it also has the highest accuracy compared with other methods.

In further study, the MOEFC method can be optimized from two aspects. On the one hand, the performance of the MOEFC method can be improved by using more advanced edge detection methods. On the other hand, the computational cost of the MOEFC method is still expensive since it is performed in the spatial domain. Therefore, an acceleration strategy needs to be developed.

ACKNOWLEDGEMENTS

This study was funded by the Key Project of Shanghai Science and technology Innovation Action (No. 20dz1201202).

REFERENCES

Carsten, S., Markus, U., Christian, W., 2008. *Machine vision algorithms and applications*, Wiley-VCH.

Dawn, S., Saxena, V., Sharma, B., 2010. Remote Sensing Image Registration Techniques: A Survey. *Proc. Int. Conf. Image and Signal Processing*, 103–112.

Fan, J., Wu, Y., Li, M., Liang, W., Cao, Y., 2018. SAR and optical image registration using nonlinear diffusion and phase congruency structural descriptor. *IEEE Trans. Geosci. Remote Sens*. 56(9), 5368–5379. doi.org/10.1109/TGRS.2018.2815523.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* 24, 381–395. doi.org/10.1145/358669.358692.

Jing, J., Liu, S., Wang, G., Zhang, W., Sun, C., 2022. Recent advances on image edge detection: A comprehensive review. *Neurocomputing* 503, 259–271.

Lan, C., Lu, W., Yu, J., Xu, Q., 2021. Deep learning algorithm for feature matching of cross modality remote sensing images. *Acta Geod. Cartogr. Sin*. 2021, 50(2), 189-202.

Li, J., Hu, Q., Ai, M., 2020. RIFT: Multi-Modal Image Matching Based on Radiation-Variation Insensitive Feature Transform. *IEEE Trans. Image Process*. 29, 3296–3310. doi.org/10.1109/TIP.2019.2959244.

Sui, H., Liu, C., Gan, Z., Jiang, Z., Chuan, X., 2022. Overview of multi-modal remote sensing image matching methods. *Acta Geod. Cartogr. Sin.* 51(9), 1848-1861.

Wang, M., Ye, Y., Zhu, B., Zhang, G., 2022. An Automatic Registration Method for Optical and SAR Images Based on Spatial Constraint and Structure Features. *Geomatics Inf. Sci. Wuhan Univ*. 47(1), 141-148.

Xiang, Y., Wang, F., You, H., 2018. OS-SIFT: A Robust SIFT-Like Algorithm for High-Resolution Optical-to-SAR Image Registration in Suburban Areas. *IEEE Trans. Geosci. Remote Sens*. 56, 3078–3090. doi.org/10.1109/TGRS.2018.2790483.

Xu, C., Sui, H.G., Li, D.R., Sun, K.M., Liu, J.Y., 2016. An automatic optical and sar image registration method using iterative multi-level and refinement model. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLI-B7, *593–600*. doi.org/10.5194/isprsarchives-XLI-B7-593-2016.

Yao, Y., Zhang, Y., Wan, Y., Liu, X., Guo, H., 2021. Heterologous Images Matching Considering Anisotropic Weighted Moment and Absolute Phase Orientation. *Geomatics Inf. Sci. Wuhan Univ*. 46(11), 1727-1736.

Ye, Y., Bruzzone, L., Shan, J., Bovolo, F., Zhu, Q., 2019. Fast and Robust Matching for Multimodal Remote Sensing Image Registration. *IEEE Trans. Geosci. Remote Sens*. 57(11), 9059–9070. doi.org/10.1109/TGRS.2019.2924684.

Ye, Y., Shan, J., Bruzzone, L., Shen, L., 2017. Robust registration of multimodal remote sensing images based on structural similarity. *IEEE Trans. Geosci. Remote Sens*. 55(5), 2941-2958. doi.org/10.1109/TGRS.2017.2656380.

Ye, Y., Tang, T., Zhu, B., Yang, C., Li, B., Hao, S., 2022. A Multiscale Framework With Unsupervised Learning for Remote Sensing Image Registration. *IEEE Trans. Geosci. Remote Sens*. 60, 1–15. doi.org/10.1109/TGRS.2022.3167644.

Zhang, Y., Yao, Y., Wan, Y., Liu, W., Yang, W., Zheng, Z., Xiao, R., 2023. Histogram of the orientation of the weighted phase descriptor for multi-modal remote sensing image matching. *ISPRS J. Photogramm. Remote Sens*. 196, 1–15. doi.org/10.1016/j.isprsjprs.2022.12.018.

Zhu, B., Zhang, J., Tang, T., Ye, Y., 2022. SFOC: a Novel Multi-Directional and Multi-Scale Structural Descriptor for Multimodal Remote Sensing Image Matching. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B2-2022, 113–120. doi.org/10.5194/isprs-archives-XLIII-B2-2022-113-2022.