A NOVEL REMOTE SENSING IMAGE REGISTRATION ALGORITHM BASED ON THE ADAPTIVE PCNN SEGMENTATION

J. F. Ge[1,2], Y. S. Zhang[1], X. J. Li[1,2,3]*, H. Li[1], Y. K. Li[1,2,3]

[1]Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China, xjli@mail,lzjtu.cn
[2]National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring, Lanzhou, China
[3]Gansu Provincial Engineering Laboratory for National Geographic State Monitoring, Lanzhou, China

Commission III and IV

KEY WORDS: Image Registration, Slime Mould Algorithm, PCNN, Remote Sensing Image, Segmentation, Feature Detection

ABSTRACT:

The appropriate feature segmentation can further improve the remote sensing image registration. The paper proposes a novel adaptive region-based registration method for remote sensing image, which combines the PCNN segmentation and feature-based method. Specifically, the parameters of PCNN are adaptively optimized by the slime mould algorithm. The reference and the input image are matched by the similar regions of PCNN segmentation, which is insensitive to the geometric and photometric changes. Then, two images are registered by the regional matching. Since the segmentation regions of the PCNN agree with the human visual system, and more stable. The proposed method achieves better registration performance. Experimental results conducted on UAV and GaoFen-2 remote sensing image pairs indicate that the proposed method outperforms the SIFT, SURF, Harris-Laplace, MSER methods.

1. INTRODUCTION

Image registration is the process of matching two images with different time, angle and sensor to the same coordinate system (Zitova et al., 2003). Remote sensing image registration is the basis of the multi-modal remote sensing information fusion. The accuracy of the registration will affect the results of subsequent image processing and analysis tasks.

The classical image registration methods can be divided into two classes: the intensity-based methods and the feature-based methods. The intensity-based methods, which are also called the template matching methods, searches for the matching relationship between the two images by the specific similarity measure templates. Commonly used similarity measures include the sum of squared difference (SSD), the normalized cross-correlation (NCC) (Mahmood et al., 2011), and the mutual information (MI) (Zhang et al., 2018). The feature-based methods calculate the parameters of the registration model by matching the robust features which detected from the two images. The feature-based registration methods include the scale invariant feature transform (SIFT) (Lowe., 2004), the speeded up robust features (SURF) (Bay et al., 2008), and the maximally stable extremal region (MSER) (Ali et al., 2016).

Feature-based methods mainly include four stages: feature detection, feature description, feature matching and geometric transformation estimation. Feature detection refers to the detection of salient image features, such as corners, intersections, spots, edges and regions (Tuytelaars et al., 2008). However, the accuracy rate of point-feature matching will be reduced due to the imaging mechanism and the impact of noise on the image in the remote sensing image registration. Compared with the feature points, the feature regions are more reliable in the remote sensing registration. Thus, extracting the similar regions of the two images will get better registration result (Okorie et al., 2019). Then these regions can be matched to improve the registration accuracy. The pulse coupled neural network (PCNN), which is an artificial neural network with physiological excitation. It can obtain irregular segmentation regions by its nonlinear activations of neurons (Johnson et al., 1999). Since its biological background and pulse coupling characteristics, the PCNN shows great potential in image processing and pattern recognition applications (Lian et al., 2021).

The paper proposes a segmentation based PCNN registration method (SBPR). Firstly, we adaptively optimize PCNN parameters by using the slime mould algorithm (SMA). And then, the similar regions from both the reference and the input images are obtained by the segmentation of PCNN. Secondly, the regions are processed by the morphological morphology and ellipse fitting algorithms. Then the feature vector of the fitting ellipse is computed by the SURF descriptor. Finally, the nearest neighbor distance ratio of different regions and the fast sample consistency (FSC) (Wu et al., 2014) are used to register two images. Compare with the classic feature-based methods, the proposed method has better registration performance, as shown by our experiments.

2. RELATED WORK

2.1 PCNN Model

The PCNN is a two-dimensional neural network model based on the principle of mammalian vision. The model structure is shown in Figure 1. The mathematical description of the standard neuron model is as follows:

* Corresponding author

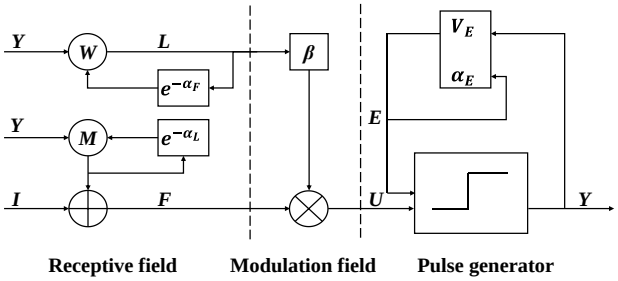


Figure 1. PCNN model.

$$F_{ij}[n] = e^{-\alpha_F} F_{ij}[n-1] + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1] + I_{ij} \quad (1)$$

$$L_{ij}[n] = e^{-\alpha_L} L_{ij}[n-1] + V_L \sum_{kl} W_{ijkl} Y_{kl}[n-1] \quad (2)$$

$$U_{ij} = F_{ij}[n](1 + \beta L_{ij}[n]) \quad (3)$$

$$Y_{ij}[n] = \begin{cases} 1 & U_{ij}[n] > E_{ij}[n] \\ 0 & otherwise \end{cases} \quad (4)$$

$$E_{ij}[n+1] = e^{-\alpha_E} E_{ij}[n] + V_E Y_{ij}[n] \quad (5)$$

The PCNN model defines each pixel of the image as a neuron. In PCNN iteration, neurons (i, j) will be affected by neighboring neurons (k, l), causing the neuron to accelerate activation or inhibit activation. The parameter I is the external stimulation. F and L represent the feeding input and the linking input, respectively. β is the linking strength. M and W represent the linking synapse weights. In an iteration, when the internal activity U_{ij} of the neuron (i, j) is greater than the dynamic threshold E_{ij}, the neuron is activated and Y_{ij} changed from 0 to 1. α_F, $\alpha_L$and α_E are time constants that used to control the decay of F, L, E, respectively. V_F, $V_L$and V_E are normalizing constants. Y is the binary matrix output of PCNN network, which can be used as the segmentation result of each iteration.

2.2 Slime Mould Algorithm

The slime mould algorithm (SMA) is a swarm intelligence optimization algorithm based on the foraging behavior of slime mould individuals (Li et al., 2020). Slime mould changes the direction of cytoplasmic flow by the concentration of food in the vein. Then, slime mould will move to the places, with the higher food concentrations. The model of the SMA is given by the following equations.

$$X(h+1) = \begin{cases} X_b(h) + v_b(W * X_A(h) - X_A(h)) & r < p \\ v_c * X(h) & otherwise \end{cases} \quad (6)$$

Where h is the iteration number. $X_b(h)$ denotes the optimal slime mould individual in the current iteration. $X_A(h)$ and $X_B(h)$ are the random positions of two individuals. W represents the weight of the slime mould. v_b and v_c are control parameters. The range of v_b is $[-a, a]$. v_c decreases from 1 to 0 linearly. r denotes a random number between [0,1]. The mathematical descriptions of the control variable p and a are shown in Eq. (7) and Eq. (8), respectively.

$$p = tanh|S(i) - DF|, \quad i \in 1,2,3, \ \ldots, \ n \quad (7)$$

$$\alpha = arctanh\ (1 - {}^{h}/_{h_{max}}) \quad (8)$$

$S(i)$ is the current individual fitness value. DF represents the best fitness value in the current iteration. h_{max} denotes the maximum number of iterations.

SMA generates the optimal path by adaptively adjusting the weights, which is consistent with the positive or negative feedback of the slime mould in the process of foraging. The mathematical description of the weight parameter W is shown in Eq. (9).

$$W(SI(i)) = \begin{cases} 1 + r * log(\frac{gF - S(i)}{gF - wF} + 1) & condition \\ 1 - r * log(\frac{gF - S(i)}{gF - wF}) & others \end{cases} \quad (9)$$

Where the range of r is [0,1]. gF is the best fitness value in the current iteration. wF represents the worst fitness value in the current iteration, and *condition* here represents the individuals that their fitness is in the top half of all population. $SI(i)$ refers to the food concentration index.

Even though the slime moulds have found the well food source position, they will still separate some individuals to explore other position with higher quality food source. Therefore, the mathematical description of the renewal position of the slime mould population is shown in Eq. (10).

$$X(h+1) = \begin{cases} rand * (ub - lb) + lb & rand < z \\ X_b(h) + v_b(W * X_A(h) - X_B(h)) & r < p \\ v_c * X(h) & r \geq p \end{cases} \quad (10)$$

The range of slime mould position is [lb, ub]. Moreover, *rand* is a random number between 0 and 1, and z is a custom parameter.

3. PROPOSED SBPR METHOD

The proposed SBPR method includes three parts: the adaptive PCNN similarity segmentation module, the segmentation regions description module, and the region matching module. Figure 2 shows the flowchart of the proposed SBPR method.

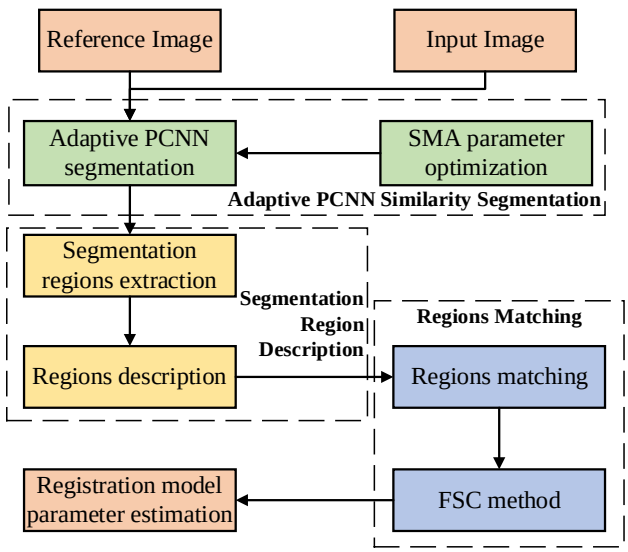


Figure 2. The architecture of the proposed SBPR method.

3.1 Adaptive PCNN Similarity Segmentation

In the traditional PCNN, its segmentation parameters are set to series of unfixed values in different image processing tasks.

However, different parameters of PCNN will lead to the different segmentation results. Furthermore, the PCNN model has many parameters which are difficult to choose. Linking strength β determines the contribution of linear linking input in the internal activities. In addition, α_F and α_E determine the falling rate of feeding input F and dynamic threshold E, respectively. These three parameters greatly affect the PCNN segmentation results. In the proposed method, the PCNN model parameters α_F, β and α_E are adaptively optimized by SMA, which to acquire the similar regions.

The fitness function consists of the root mean square error (RMSE) of the matched region and the correct matching rate (CMR). The mathematical description of the fitness function is shown in Eq. (11).

$$fitness = (\gamma_1 * e^{-RMSE} + \gamma_2 * CMR) * N_{cor} \quad (11)$$

Where N_{cor} denotes the number of correct matched region. γ_1 and γ_2 are fitness weights and $\gamma_1 + \gamma_2 = 1$.

The optimization process of the slime mould algorithm for PCNN model is as follows:

(1) Initialize the position of each slime mould by adjusting the PCNN model parameter α_F, β and α_E.
(2) Substitute the position of slime mould into PCNN model for segmentation. Then the irregular segmentation regions are matched to calculate the fitness value.
(3) Update the slime moule position according to the food searching rules. Then, the fitness function is updated to the optimal fitness value.
(4) Let iteration times $t = t + 1$, and returns to step 3, if $t < t_{max}$.
(5) Stop the iteration when $t \geq t_{max}$. And obtain the optimal parameters of the PCNN similarity segmentation.

3.2 Segmentation Region Description

The segmentation region of the PCNN model is irregular, which is not convenient to describe its features. In order to obtain the accurate description, the regions are processed by the morphological morphology and ellipse fitting algorithms. Then the feature vector of the fitting ellipse is computed by SURF descriptor.

The centroid of the region is calculated by the geometric zero moment and the geometric first moment of the region. The centroid of the irregular region is taken as the centre of the fitting ellipse region. The mathematical description is as follows:

$$x_n = \frac{\sum xR_n(x,y)}{\sum R_n(x,y)} \quad (12)$$

$$y_n = \frac{\sum yR_n(x,y)}{\sum R_n(x,y)} \quad (13)$$

Where x_n and y_n are the centroid of region R_n.

The second moment of the irregular region is used to determine the major axis a, the minor axis b and the direction of the major axis θ. The mathematical description of the second-order moment is as follows:

$$U = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} \quad (14)$$

$$a = \sqrt{\frac{(\mu_{20}+\mu_{02})+\sqrt{(\mu_{20}+\mu_{02})^2+4\mu_{11}^2}}{2\sum R_n(x,y)}} \quad (15)$$

$$b = \sqrt{\frac{(\mu_{20}+\mu_{02})-\sqrt{(\mu_{20}+\mu_{02})^2+4\mu_{11}^2}}{2\sum R_n(x,y)}} \quad (16)$$

$$\theta = \frac{1}{2} arctan\left(\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right) \quad (17)$$

Where U is the second moment of region R_n. $\mu_{20} = \sum(x - x_n)^2 R_n(x,y)$, $\mu_{02} = \sum(y - y_n)^2 R_n(x,y)$, $\mu_{20} = \sum(x - x_n)(y - y_n) R_n(x,y)$.

The SURF descriptors are calculated by the centroid (x_n, y_n), the major axes a, the minor axes b and the direction of the major axis θ. The centre of the feature description region is (x_n, y_n). The side length and the main direction of the feature description region are $\sqrt{a^2 + b^2}$ and θ, respectively. The feature description region is divided into 16 sub regions. The sum of the horizontal and vertical components of the harr wavelet response value ($\sum dx$ and $\sum dy$)and the sum of the horizontal and vertical components of the absolute values of the harr wavelet response value ($\sum |dx|$ and $\sum |dy|$) are calculated in each sub region. Additionally, the feature vectors are composed by $\sum dx$, $\sum dy$, $\sum |dx|$ and $\sum |dy|$.

3.3 Region Matching

In the paper, the similarities between the two regions are measured by the nearest neighbor distance ratio (NNDR). NNDR calculates the ratio of the nearest neighbor distance to the secondary neighbor distance of the descriptors. The feature vectors are considered as the matching pairs, if the ratio exceeds the threshold. In the proposed method, the FSC algorithm is also used to eliminate the mismatched pairs to reduce the influence of mismatched pairs.

4. EXPERIMENTAL RESULTS

Experiments are carried out on two high-resolution remote sensing datasets, i.e., the Mavic2 UAV dataset and the GaoFen-2 remote sensing dataset to exam the effectiveness of the proposed SBPR method.

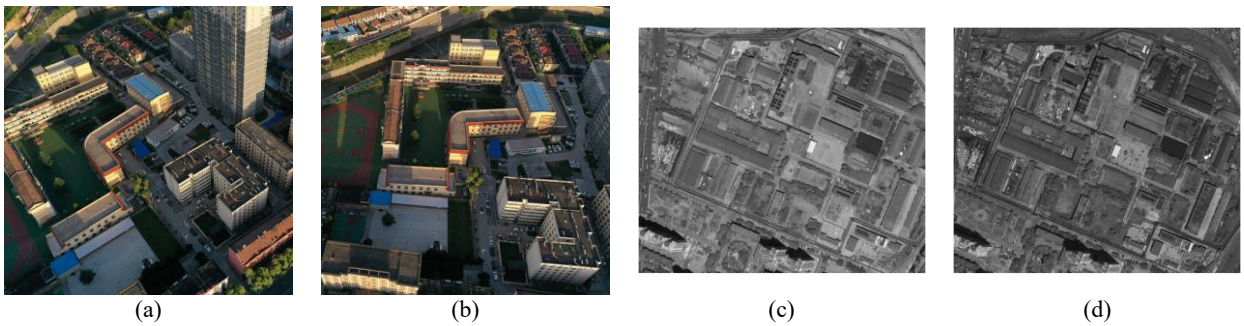
4.1 Datasets for Comparison

The first dataset was collected by Mavic2 UAV in Baoji city, China. The size of all images in the dataset is 700×700 pixels. The spatial resolution of the images is 0.1m×0.1m. And the images in the dataset have great visual angle changes.

The other dataset is the urban area of LanZhou city in China, which is obtained by the GF-2 remote sensing satellite sensor. The images size of the dataset is 700×600 pixels. The spatial resolution of the images is 1m×1m. The images of the two datasets are shown in Figure 3.

4.2 Quality Assessment Criteria

The effectiveness of the proposed SBPR method is examined by the RMSE evaluation criterion, the CMR evaluation criterion and the error statistics of matched inliers (ESMI) evaluation criterion. (Fan et al., 2016, Ma et al., 2016, Yu et al., 2019). The ideal value for RMSE is 0. The mathematical description of the RMSE is shown in Eq. (18).



(a) (b) (c) (d)

Figure 3. Registration datasets. (a) UAV reference image (b) UAV input image (c) GF-2 reference image (d) GF-2 input image

$$RMSE = \sqrt{\frac{1}{NCM}\sum_{i=1}^{NCM}[(x_1^i - x_2^i)^2 + (y_1^i - y_2^i)^2]^2} \quad (18)$$

Where *NCM* represents the number of initial matching pairs, (x_1^i, y_1^i) are gravity centre coordinate of the match region in the reference image. Moreover, (x_2^i, y_2^i) denotes the gravity centre coordinates of the match region after affine transformation.

CMR ideal value is 1. The mathematical description of CMR is as follows:

$$CMR = \frac{N_{cor}}{NCM} \quad (19)$$

Where N_{cor} denotes the number of the correct matched regions. After transformation, the region pairs, in which the gravity centre distance is less than 2 pixels, are selected as the correct matched result.

The ideal value for ESMI is 0, which is defined as follows:

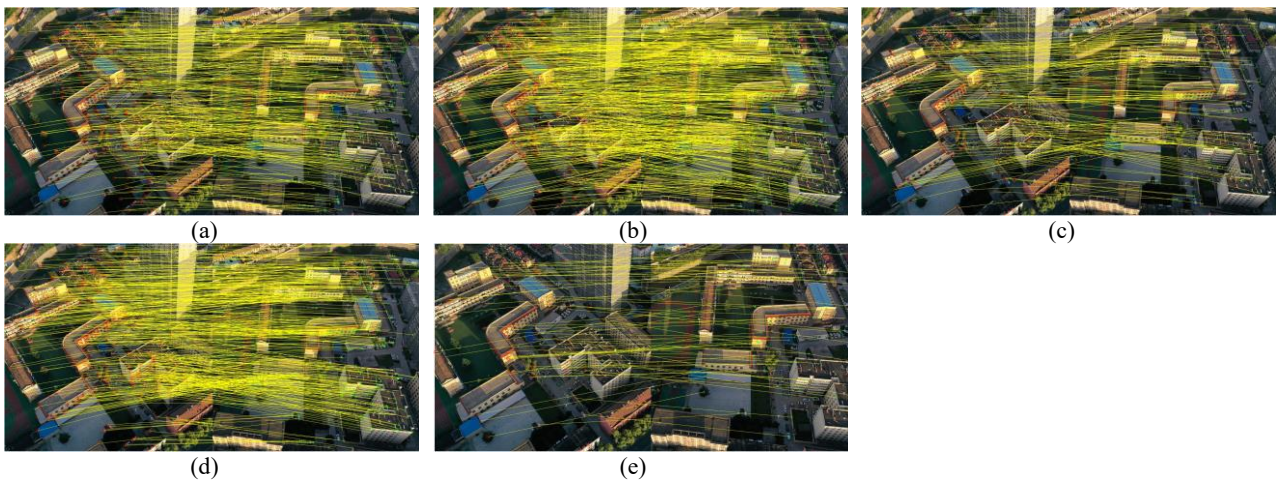
$$Var_x = \frac{1}{NCM}\sum_{i=1}^{NCM}\left[(x_1^i - x_2^i) - \frac{1}{NCM}\sum_{i=1}^{NCM}(x_1^i - x_2^i)\right]^2 \quad (20)$$

$$Var_y = \frac{1}{NCM}\sum_{i=1}^{NCM}\left[(y_1^i - y_2^i) - \frac{1}{NCM}\sum_{i=1}^{NCM}(y_1^i - y_2^i)\right]^2 \quad (21)$$

Where Var_x and Var_x are the variances of the matching error in the *x* and *y* directions, respectively.

4.3 Experimental Results and Analysis

In order to examine the performance of the SBPR method, we choose SIFT, SURF, Harris-Laplace (MIKOLAJCZYK et al. 2004) and MSER methods for comparative experiments. The SIFT, SURF and Harris-Laplace belongs to point-feature-based methods. And the MSER belongs to the region-feature-based methods. Moreover, the MSER methods uses the same feature description as SBPR methods. It should be noted that all five methods exploit the nearest neighbor distance ratio and FSC methods to perform feature matching. Figure 4 and Figure 5 show the matching results of dataset UAV and dataset GF-2, respectively. Figure 6 shows the Checkerboard mosaicked images of dataset UAV and dataset GF-2. Table 1 and Table 2 show the objective evaluation of the two datasets matching results respectively.

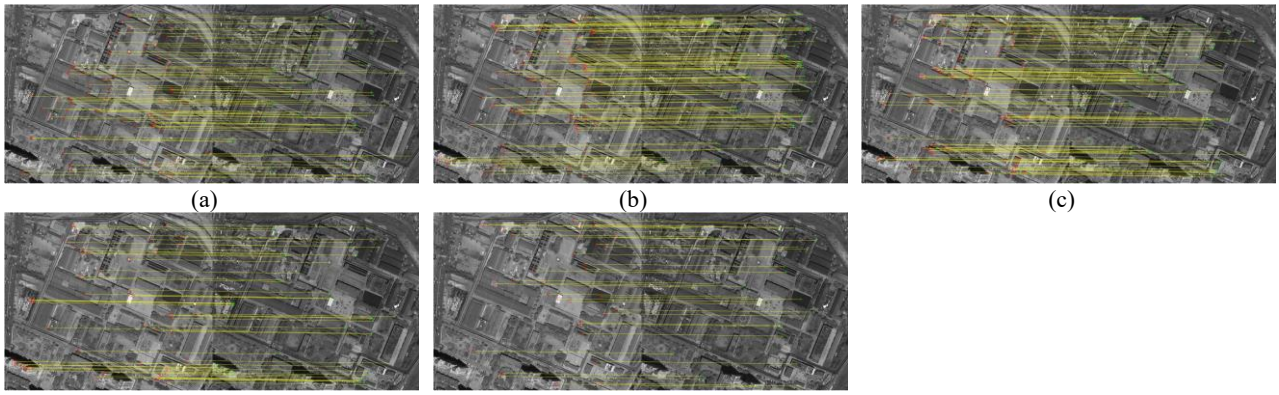


(a) (b) (c) (d) (e)

Figure 4. Matching results of UAV dataset. (a) SBPR (b) SIFT (c) SURF (d) Harris-Laplace (e) MSER

Methods	N_{cor}	NCM	CMR	RMSE	Var_x	Var_y
SBPR	311	333	0.9389	0.6635	0.1269	0.0975
SIFT	505	562	0.8956	0.7384	0.1426	0.1070
SURF	162	188	0.8617	0.8442	0.1344	0.1651
Harris-Laplace	426	505	0.8436	0.7574	0.1861	0.1171
MSER	71	87	0.8506	0.8100	0.2613	0.1006

Table 1. Comparison results for the UAV dataset

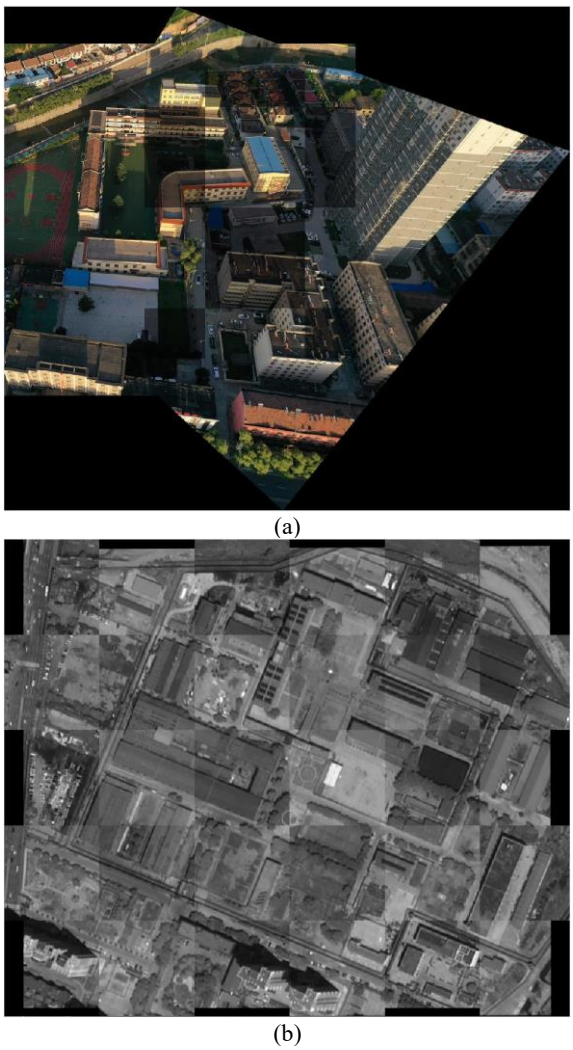


(a) (b) (c) (d) (e)

Figure 5. Matching results of GF-2 dataset. (a) SBPR (b) SIFT (c) SURF (d) Harris-Laplace (e) MSER

Methods	N_{cor}	NCM	CMR	RMSE	Var_x	Var_y
SBPR	76	92	0.8261	0.9072	0.1782	0.1087
SIFT	97	135	0.7185	1.0094	0.1847	0.1692
SURF	96	120	0.8000	0.9890	0.3236	0.1153
Harris-Laplace	56	73	0.7671	1.1129	0.3634	0.1247
MSER	33	62	0.5323	1.0665	0.2822	0.2340

Table 2. Comparison results for the GF-2 dataset



(a) (b)

Figure 6. The checkerboard mosaicked images of datasets. (a) UAV dataset (b) GF-2 dataset

The Figure 4 and Figure 5 show the matching results of UAV dataset and GF-2 dataset, respectively. It can be seen from the Figure 4 and Figure 5 that the MSER method has obtained the least matching point pairs. Because MSER describes the Gray-level stable region as feature region, which is easy to be affected by noise and Gray-level change. The SURF method has fewer matching pairs in the UAV dataset, which is caused by the large geometric distortion in the UAV dataset. The SIFT method and SBPR method achieve better matching performance in both datasets.

The Table 1 and Table 2 are the objective evaluation results of UAV dataset and GF-2 dataset, respectively. It can be seen from Table 1 and Table 2 that although the SBPR method and the MSER method adopt the same feature description, the N_{cor}, RMSE and ESMI indexes of the SBPR method are obviously superior to MSER method. Compared with SIFT method, the NCM index of SBPR method is worse. The reason is that the SBPR method uses the segmentation region as the matching feature. It can be observed that the CMR, RMSE and ESMI indexes perform best when compared with the other methods, which demonstrates that the proposed method is effective in the image registration task. In addition, from the checkerboard mosaicked images in Figure 6, it can be clearly seen that the features of the reference image and the input image are completely overlapped.

5. CONCLUSION

In the paper, a novel adaptive region-based registration method is proposed. Firstly, the PCNN model parameters are adaptively optimized by SMA. Therefore, the PCNN model could obtain more similar regions between different remote sensing images. Additionally, the regions are processed by the morphological morphology and ellipse fitting algorithms. Then the feature vector of the fitting ellipse is computed by SURF descriptor. Finally, the nearest neighbor distance ratio and the FSC are used to register two images. Through two groups of comparative experiments, the SBPR method has shown the best registration

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W2-2022-17-2022 | © Author(s) 2022. CC BY 4.0 License.

precision than the SIFT, SURF, Harris-Laplace and MSER methods.

Although the propose method has the well registration precision, the matching points are less than some feature-based methods. Thus, how to combine the region features and the point features will be our subsequent work. The segmentation performance will be improved in the future work

ACKNOWLEDGEMENTS

This paper is jointly supported by National Natural Science Foundation of China (No.41861055), China Postdoctoral Science Foundation (No.2019M653795) and lzjtu EP Program (No.201806).

REFERENCES

Ali, F., Khan, S U., Mahmudi, M Z., et al. 2016. A comparison of FAST, SURF, Eigen, Harris, and MSER features. *International Journal of Computer Engineering and Information Technology*, 8(6): 100.

Bay, H., Ess, A., Tuytelaars, T., et al. 2008. Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3): 346-359.

Fan, J., Wu, Y., Li, M., et al. 2016. SAR image registration using multiscale image patch features with sparse representation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(4): 1483-1493.

Johnson, J L., Padgett, M L. 1999. PCNN models and applications. *IEEE transactions on neural networks*, 10(3): 480-498.

Lian, J., Yang, Z., Liu, J., et al. 2021. An overview of image segmentation based on pulse-coupled neural network. *Archives of Computational Methods in Engineering*, 28(2): 387-403.

Lowe, D G. 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2): 91-110.

Ma, W., Wen, Z., Wu, Y., et al. 2016. Remote sensing image registration with modified SIFT and enhanced feature matching. *IEEE Geoscience and Remote Sensing Letters*, 14(1): 3-7.

Mahmood, A., Khan, S. 2011. Correlation-coefficient-based fast template matching through partial elimination. *IEEE Transactions on image processing*, 21(4): 2099-2108.

MIKOLAJCZYK, K., SCHMID, C. 2004. Scale and affine invariant interest point Detectors. *International Journal of Computer Vision*, 60 (1): 63-86.

Okorie, A., Makrogiannis, S. 2019. Region-based image registration for remote sensing imagery. *Computer Vision and Image Understanding*, 189, 102825.

Tuytelaars, T., Mikolajczyk, K. 2008. Local invariant feature detectors: a survey. *Foundations and trends in computer graphics and vision*, 3(3): 177-280.

Wu, Y., Ma, W., Gong, M., et al. 2014. A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geoscience and Remote Sensing Letters*, 12(1): 43-47.

Yu, Q., Zhou, S., Jiang, Y., et al. 2019. High-performance SAR image matching using improved SIFT framework based on rolling guidance filter and ROEWA-powered feature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(3): 920-933.

Zhang, J., Zareapoor, M., He, X., et al. 2018. Mutual information based multi-modal remote sensing image registration using adaptive feature weight. *Remote Sensing Letters*, 9(7): 646-655.

Zitova, B., Flusser, J. 2003. Image registration methods: a survey. *Image and vision computing*, 21(11): 977-1000.