A NOVEL ROBUST FEATURE DESCRIPTOR FOR MULTI-SOURCE REMOTE SENSING IMAGE REGISTRATION

Song Cui[1], *Yanfei Zhong*[1,]*, *Ailong Ma*[1], *Liangpei Zhang*[1]

[1]State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, P. R. China
*Corresponding author E-mail: zhongyanfei@whu.edu.cn Phone: 86-27-68779969

ABSTRACT

Non-linear radiation difference (NRD) will lead to the corresponding features cannot be mapped one by one, so the traditional image feature matching methods based on intensity or gradient fail to be directly applied to the multi-source remote sensing image registration. In this paper, a new robust feature descriptor is proposed, which has the invariance of radiation, scale and rotation. The nonlinear diffusion function which is insensitive to the radiation difference is used to construct the scale space so that the descriptors can be used in images with different resolutions. A pixel-by-pixel local phase congruency algorithm is used to extract the corresponding points, and then the features are described by means of rotation invariance description. Feature matching is completed based on feature vector constructed by the descriptor, thus to realize image registration. In the experimental part, three kinds of multi-source remote sensing images with large radiation differences were used to test the descriptors. The results showed that the proposed method effectively extracted the corresponding features, and achieved the best effect in the quantitative evaluation of image registration.

Index Terms—Image registration, nonlinear radiation distortions (NRD), phase consistency, structural descriptor

1. INTRODUCTION

Image registration is an essential and fundamental task in remote sensing, which is aimed to match two or more images obtained from different imaging conditions, different sensors or different perspectives [1]. The complementary information between multi-source remote sensing images can effectively improve the interpretation effect. A large number of non-linear radiation distortions (NRD) present in multi-source remote sensing images as their different physical imaging mechanism, thus traditional method is failed to deal with such distortions, which becomes difficult to register those images.

In view of multi-source remote sensing images registration, researchers have proposed various descriptors based on structure information whose sensitivity to NRD is less than gradient. Shechtman and Irani [2] introduce a local self-similarity (LSS) descriptor that captures the internal geometric layout of an image based on a logarithmic polar coordinate grid. The histogram of orientated phase congruency (HOPC) [3] algorithm expends the phase congruency mode to include digital and directional information. Phase congruency based structural descriptor (PCSD) [4] is also based on PC without rotational distortions. Radiation-invariance feature transformation (RIFT) [5] is a rotation invariance method which is suitable for multi-source remote images without scale invariant. The Euclidean distance of multi-scale phase consistency (MS-PC) descriptors, which is more robust to the radiation differences between images, is used as similarity metric to achieve correspondences [6].

In this paper, a new robust feature descriptor is proposed for feature matching to the registration of multi-source remote sensing images. The contribution of this paper can be summarized as follows.

(1) The nonlinear diffusion function which is insensitive to the radiation difference is used to construct the scale space so that the descriptors can be used in images with different resolutions. Compare with the Gaussian function, it preserves the edges and details better while suppressing noise.

(2) A pixel-by-pixel local phase congruency algorithm is used to establish correspondences between images with NRD and extract the corresponding points, so that the feature of those images are mapped into the same space which can be easily described.

(3) A locally rotationally invariant coordinate system was used for calculation of the feature descriptor. the sample points neighbor a key point are statistically calculated through a constantly changing local coordinate system, which itself realizes the rotation invariance without the need for the dominant direction.

2. FEATURE DESCRIPTOR CONSTRUCTION

In this section, the proposed descriptor method is elaborated. Fig. 1 presents a flowchart of the proposed feature descriptor construction. Initially, the nonlinear diffusion scale space is established. Secondly, the pixel-by-pixel local phase congruency is computed. Finally, the locally rotationally invariant coordinate system is constructed.

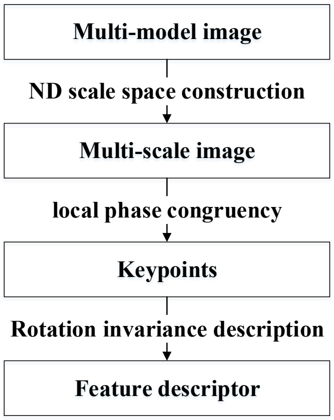


Fig. 1. Flowchart of the proposed descriptor construction.

2.1. Scale invariance based on nonlinear diffusion scale space construction

To capture more salient feature points, in the ND scale space construction, nonlinear diffusion function is adopted to replace Gaussian function which is usually used for SIFT detector. Compared with Gaussian function, nonlinear diffusion function has proven to preserve the edges and details better while suppressing noise [7]. The nonlinear diffusion function can be expressed as,

$$\begin{aligned}\frac{\partial f(x,y)}{\partial t} &= f_t = div\left(c(x,y,t)\nabla f\right) \\ &= c(x,y,t)\Delta f + \nabla c \bullet \nabla f\end{aligned} \quad (1)$$

Where t is a scale parameter, div is a bifurcation operator, ∇ and Δ are gradient and Laplace operators, respectively. $c(x,y,t)$ is the diffusion coefficient. In some special case that the nonlinear diffusion is assumed to be isotropic, in other words, the $c(x,y,t)$ is a constant, the above formula is equivalent to the Gaussian function.

The diffusion coefficient $c(x,y,t)$ is generally regarded as $|\nabla f|$,a nonlinear function of image gradient. In order to preserve the edge of the image while ensuring a faster diffusion rate, the following formula is selected as $c(x,y,t)$,

$$c = \frac{1}{1+\left(\frac{|\nabla f|}{K}\right)^2} \quad (2)$$

The parameter K, which controls which edges have to be kept or cancelled, is the factor regulatory the degree of diffusion. The greater K value chosen, the less edge information will be preserved.

In order to apply the algorithm to images with different resolutions, the scale space is discretized into a series of octaves and sublevels. By using the original image as an initial condition, the scale values are calculated from the following expression,

$$\sigma_i(o,s) = \sigma_o 2^{o+\frac{s}{S}}, o \in [0,...,O-1], s \in [0,...S+2], i \in [0,...W-1] \quad (3)$$

where $\sigma 0$ is the base scale level, and W is the total number of the smoothed images; o and s are the index of octave O and sublevel S, respectively.

2.2. Radiation invariance relied on local phase congruency(PC)

The Phase information which is obtained by the Fourier transform of image has high invariance to image contrast, illumination, scale, and other changes. Further, the degree of consistency of the local phase information at different angles is called PC measure.

Defines a two-dimensional image as $I(x,y)$, the singular symmetry part of wavelet transform $O_{so}(x,y)$ and even symmetrical parts $E_{so}(x,y)$ convolve image $I(x,y)$ to obtain,. In scale 0 and orientation S, the amplitude part and the phase part of image $I(x,y)$ can be expressed as,

$$A_{so}(x,y) = \sqrt{E_{so}(x,y)^2, O_{so}(x,y)^2} \quad (4)$$

$$\phi_{so}(x,y) = \arctan(O_{so}(x,y) / E_{so}(x,y)) \quad (5)$$

all scales 0 and all orientation S are considered, the calculation results of the two-dimensional phase congruency are calculated as follows,

$$PC(x,y) = \frac{\sum_s \sum_o w_o(x,y) \lfloor A_{so}(x,y)\Delta\Phi_{so}(x,y) - T \rfloor}{\sum_s \sum_o A_{so}(x,y) + \xi} \quad (6)$$

where $w_o(x,y)$ is a weight function; ξ is a constant with a very small number; $\lfloor \ \rfloor$ action is to prevent negative values, which means that when the value is negative, its result is 0. $A_{so}(x,y)\Delta\Phi_{so}(x,y)$ is a phase deviation function.

Before calculation of the minimum and maximum moments, three intermediate quantities are computed,

$$a = \sum_o \left(PC(\theta_o)\cos(\theta_o)\right)^2 \quad (7)$$

$$b = 2\sum_o \left(PC(\theta_o)\cos(\theta_o)\right)\left(PC(\theta_o)\sin(\theta_o)\right) \quad (8)$$

$$c = \sum_o \left(PC(\theta_o)\sin(\theta_o)\right)^2 \quad (9)$$

Then, the principal axis ψ, minimum moment M_ψ, and maximum moment m_ψ are given by,

$$\psi = \frac{1}{2}\arctan\left(\frac{b}{a-c}\right) \quad (10)$$

$$M_\psi = \frac{1}{2}\left(c + a + \sqrt{b^2 + (a-c)^2}\right) \quad (11)$$

$$m_\psi = \frac{1}{2}\left(c + a - \sqrt{b^2 (a-c)^2}\right) \quad (12)$$

The minimum moment m_ψ is equivalent to a measure for corner detector. In other words, if the value of m_ψ at a point is large, the point is most likely to be a corner feature; and the maximum moment M_ψ is the edge map of an image, which can be used for edge feature detection.

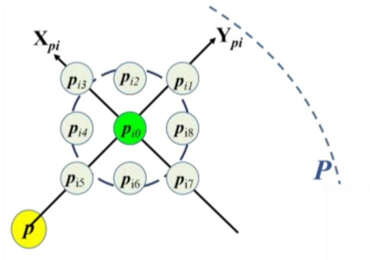


Fig. 2. The locally rotation invariant coordinate system.

2.3. Rotation invariant depended on locally rotationally invariant coordinate system

We first construct a rotationally invariant coordinate system for each key point. Suppose p_{i0} is a point in one support region of the key point p, the line connecting p and p_{i0} is set as the Y-axis, and then construct a local Cartesian coordinate system XY, in which it can be observed that the pixels in the original neighborhood of the sample point pi remain the same in the rotated neighborhood of the corresponding sample point p_i. Therefore, the local coordinate system, as shown in the Fig.2, is rotate-invariant.

For each sample point p_{id}, its difference of local orientation scale phase congruency can be computed in locally rotationally invariant coordinate system. The calculation formula is as follows,

$$Dx(X_i) = I(X_i^1) - I(X_i^5), \quad (13)$$

$$Dy(X_i) = I(X_i^3) - I(X_i^7), \quad (14)$$

where X_{ji}, $j = 1,3,5,7$ are X_i's neighboring points along the x-axis and y-axis in the local XY-coordinate system, and $I(X_{ji})$ stands for the intensity at X_{ji}. Then, the Difference of Local orientation scale phase congruency $D(X_i)$ and orientation $\theta(X_i)$ can be computed as,

$$D(X_i) = \sqrt{Dx(X_i)^2 + Dy(X_i)^2}, \quad (15)$$

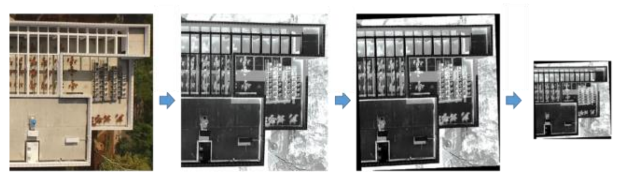
$$\theta(X_i) = \tan^{-1}(Dy(X_i) / Dx(X_i)). \quad (16)$$

3. EXPERIMENT AND ANALYSIS

In this section, the proposed algorithm is tested using remote sensing images of four different modes with intensity, scale, and rotation distortions. The experimental results show that our descriptor can make the image features with nonlinear mapping correspond one to one. By comparing with the existing methods, the proposed algorithm has the highest accuracy, the most robustness and the most versatility

3.1. Data sets

Experimental images can be divided into two categories: simulated data and real data sets. The simulated image is produced by artificially adding nonlinear intensity stretch, scale transformation and rotation transformation to an optical image, as shown in Fig. 3.



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

Fig. 3. Simulation data. (a) Original image (b) nonlinear intensity stretch (c) 5 degrees of rotation (d) 2 reduced sampling with resolution.

Real data sets contain four different modes of multi-source remote sensing images, including optical images, SAR images, LiDAR intensity images and infrared images. In these cases, the sensors and positions used for acquiring the images are not the same, therefore, the intrinsic and extrinsic parameters are different. In addition to this, the tone mapping changes in a non-linear way due to changes in illumination conditions and its different modality. To quantitatively compare different solutions on this dataset, the ground truth is obtained in the same way as in paper [8], which has been strictly marked and checked manually

TABLE 1. Descriptions of data sets used in experiments

No.	Sensor	Platform	Surface type	Size	GSD
1	Optical	Satellite	Urban	1500*1500	2m
	SAR	Satellite	Urban	1500*1500	2m
2	Optical	UAV	Vegetation	800*800	1m
	LiDAR	UAV	Vegetation	400*400	2m
3	Optical	Satellite	farmland	1000*1000	10m
	Infrared	Satellite	farmland	2000*2000	20m

3.2. Experimental results and analysis

The matched feature points for simulated data are shown in Fig. 4. The proposed algorithm can match a large number of corresponding features, considering both quantity and distribution.

The feature points of correct matching of real data are shown in Fig. 5. It can be seen that the number of matching points is large and the distribution is even, which satisfies the condition of the transformation model in image registration. The matching performance of the proposed descriptor on the image pairs with NRD is far superior to the current popular feature matching methods. There are two main reasons: (1) In feature detection stage, the PC map is established instead of image intensity, while considering both quantity and distribution of the feature, which lays a foundation for subsequent matching. (2) In the feature description stage, the sample points neighbor a key point are statistically calculated through a constantly changing local coordinate system, which itself realizes the rotation invariance without the need for the dominant direction.

The image registration results of real data are shown in Fig. 6 by the checkerboard method. As can be seen from the figure, the error between the image pairs after registration is less than one pixel.

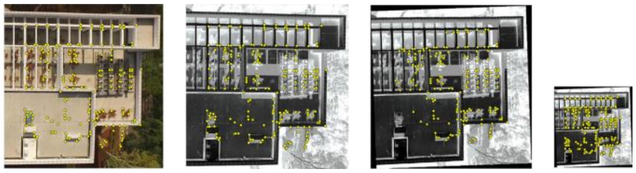
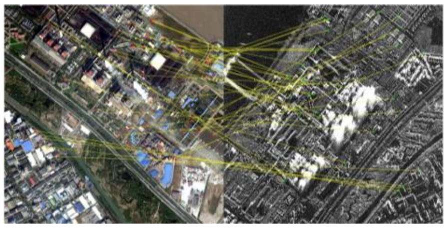
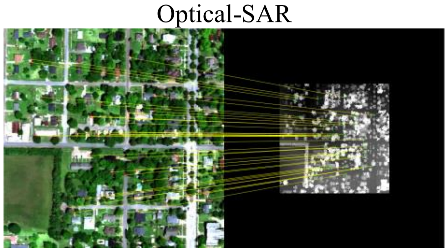


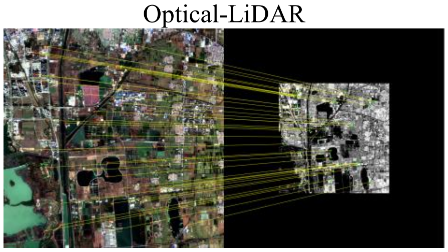
Fig. 4. Matched feature points for simulated data.



Optical-SAR

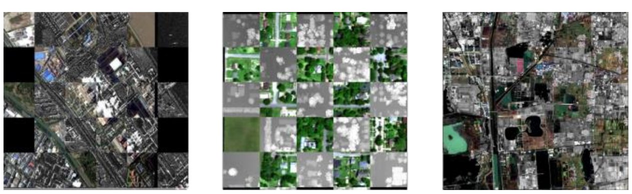


Optical-LiDAR



Optical-Infrared

Fig. 5. Feature points of correct matching of real data.



Optical-SAR Optical-LiDAR Optical-Infrared

Fig. 6. Image registration results of real data.

4. CONCLUSION

This paper proposes a new robust feature descriptor which has the invariance of radiation, scale and rotation. First of all, the ND function which is insensitive to the radiation difference is used to construct the scale space so that the descriptors can be used in images with different resolutions. Secondly, the pixel-by-pixel local phase congruency algorithm is used to extract the corresponding points, and then the features are described by means of rotation invariance description. Finally, feature matching is completed based on feature vector constructed by the descriptor, thus to realize image registration. In the experimental part, four kinds of multi-source remote sensing images with large radiation differences were used to test the descriptors. The results showed that the proposed method effectively extracted the corresponding features, and achieved the best effect in the evaluation of image registration. The matching performance via the proposed descriptor is far superior to the current feature matching methods. In future work, the proposed descriptor is planned to be compared quantitatively with some state-of-art algorithms, and extended to more modal images

5. ACKNOWLEDGEMENTS

This work was supported by National Natural Science Foundation of China under Grant Nos. 41622107, 41771385 and 41371344, National Key Research and Development Program of China under Grant No. 2017YFB0504202, and Natural Science Foundation of Hubei Province in China under Grant No. 2016CFA029.

6. REFERENCES

[1] B. Zitová and J. Flusser, "Image registration methods: a survey," Image & Vision Computing, vol. 21, pp. 977-1000, 2003.

[2] E. Shechtman and M. Irani, "Matching Local Self-Similarities across Images and Videos," in Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, 2007, pp. 1-8.

[3] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust Registration of Multimodal Remote Sensing Images Based on Structural Similarity," IEEE Transactions on Geoscience & Remote Sensing, vol. 55, pp. 2941-2958, 2017.

[4] J. Li, Q. Hu, and M. Ai, "RIFT: Multi-modal Image Matching Based on Radiation-invariant Feature Transform," 2018.

[5] J. Fan, Y. Wu, M. Li, W. Liang, and Y. Cao, "SAR and Optical Image Registration Using Nonlinear Diffusion and Phase Congruency Structural Descriptor," IEEE Transactions on Geoscience and Remote Sensing, pp. 1-12, 2018.

[6] S. Cui, and Y. Zhong, "Multi-Modal Remote Sensing Image Registration Based on Multi-Scale Phase Congruency." pp. 1-5.

[7] J. Fan, W. Yan, W. Fan, Z. Qiang, G. Liao, and L. Ming, "SAR Image Registration Using Phase Congruency and Nonlinear Diffusion-Based SIFT," IEEE Geoscience & Remote Sensing Letters, vol. 12, pp. 562-566, 2014.

[8] G. Maoguo, Z. Shengmeng, J. Licheng, T. Dayong, and W. Shuang, "A Novel Coarse-to-Fine Scheme for Automatic Image Registration Based on SIFT and Mutual Information," IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 7, pp. 4328-4338, 2014.