Fine Registration for VHR Images Based on Superpixel Registration-Noise Estimation

Xianzhang Zhu, Hui Cao, Yongjun Zhang, Kai Tan, and Xiao Ling

***Abstract*— Local nonlinear geometric distortion is problematic in the registration of very high-resolution (VHR) images. In the standard registration approach, the precision of control points generated from salient feature matching cannot be guaranteed. This letter introduces a novel superpixel registration-noise (RN) estimation method based on a two-step fine registration technique that can be estimate and mitigate the local residual misalignments in VHR images. The first step employs superpixel sparse representation and multiple displacement analysis to estimate RN information of the preregistered image. The second step optimizes the control points obtained in preregistration by combining the RN information and gross error information, and finally fine registers the input image by employing local rectification. The experiments using two data sets generated from Chinese GF2, GF1, and ZY3 satellites are discussed in this letter, and the promising results verify the effectiveness of the proposed new method.**

***Index Terms*— Image registration, local rectification, registration noise (RN), sparse representation, superpixel segmentation.**

I. Introduction

IMAGE registration is the process of overlaying two or more images of the same geographical area taken at different times, viewpoints, and modalities or by different sensors [1]. Accurate registration of images is a prerequisite of many remote sensing applications, such as nonsupervision change detection, image mosaicking, and image fusion. During the past few decades, numerous registration methods have been proposed. These methods can be categorized into two categories, area based and feature based in terms of image matching. In the case of area-based methods, the control points are obtained by calculating the maximum correlation between the image subsets. The key technique of this kind of methods is the selection of a suitable similarity measure, such as the widely used normalized correlation coefficient (NCC) [2] and mutual information [3]. In the feature-based category, the salient features are first extracted from the images and then feature matching is performed using feature correlations. Scale-invariant feature transform (SIFT) [4] and Harris points [5] are among the most popular approaches for salient feature extraction. A variety of registration algorithms [6]–[8] have been proposed to improve the standard SIFT and Harris approach. Paul and Pati [6] represented the modified uniform robust SIFT (M UR-SIFT) to register the remote sensing optical images. Based on the combination of the two categories, Ma *et al*. [9] proposed a two-step nonrigid automatic registration scheme by using SIFT and NCC. In general, feature-based method is more effective than the area-based method for the registration of very high-resolution (VHR) images because of its lower computational complexity and strong robustness. In addition, methods based on image segmentation (IS) technique have been developed in recent years, such as combination of IS and SIFT [10], histogram-based IS [11], and integration of global SRTM and segmentation (ISS) [12]. These methods utilize IS as a spatial constraint to remove regions that are not appropriate for image matching and to generate more reliable matching results.

However, the registration results of the above methods are all subject to the accuracy of control points. The terrains covered of VHR images are usually complicated, which makes nonlinear geometric distortion a ubiquitous problem. Thus, it is difficult to eliminate control point errors completely, even with the support of gross error elimination technique. This mismatch phenomenon significantly affects the creation of converting model between the input image and the reference image, which leads to local distortion of the registration results as well as unexpected effects in subsequent remote sensing applications. In the literature, some attempts have been made to analysis registration noise (RN). Bruzzone and Cossu [13] proposed a change vector analysis-based approach to reduce the effects of RN in unsupervised change detection. In [14] and [15], two algorithms (multiscale analysis based and edge based) were presented to estimate the distribution of RN. Han *et al*. [16] studied the reduction of local residual misalignment, and then proposed a segmentation-based fine registration (SBFR) approach. In their method, multitemporal images were assumed to be standardly registered and a piecewise linear function was used to achieve accurate and precise geometric alignment [16]. While extracting the control point pairs, object representative points were defined as the centroids of segments computed on the reference image. Corresponding points in the input image were located by the residual local misalignment information. However, there are few apparent and smooth geometric features in the image. This problem will adversely affect the performance of local residual misalignment estimation, and thus the accuracy of object representative point pairs.

To address the aforementioned problems, an improved fine registration method based on superpixel RN estimation (SRNE) is proposed for standardly registered (also called preregistered in this letter) VHR multitemporal and/or multisensor images. Superpixel sparse representation (SSR) is

Manuscript received April 11, 2018; revised May 21, 2018; accepted June 14, 2018. Date of publication July 16, 2018; date of current version September 26, 2018. This work was supported by the National Natural Science Foundation of China under Grant 41571434 and Grant 41322010. *(Corresponding author: Hui Cao.)*

X. Zhu, H. Cao, Y. Zhang, and X. Ling are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China (e-mail: zxzorigin@whu.edu.cn; huicao@whu.edu.cn; zhangyj@whu.edu.cn; zyl_whu@163.com).

K. Tan is with the Wuhan Research Institute of HUAWEI Technology Co., Ltd., Wuhan 430200, China (e-mail: kai_tan@whu.edu.cn).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2018.2849696

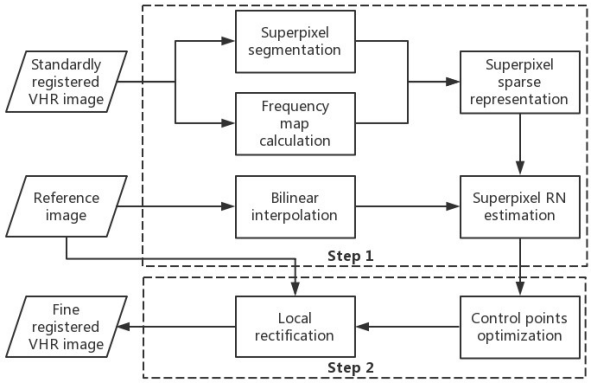

Fig. 1. Flowchart of the proposed fine registration approach.

utilized for SRNE in the input image, and then the control points are optimized to apply the local transformation [17], [18]. The main contributions of the proposed approach are twofold: 1) sparse superpixels are generated by frequency filtering so that more reliable RN estimation results can be obtained and 2) accurate control points are obtained by optimization of the preregistration matching results, where the gross error information and correlation coefficients between the input image and the reference image are fully considered.

Section II describes the methodology of the SRNE-based fine registration process; Section III discusses the experimental results; and finally, conclusions are given in Section IV.

II. Proposed Approach

The purpose of fine registration is to mitigate local residual misalignment between the standardly registered input image and the reference image. Spatial correlation-based local displacement detection between two images is very similar to interactive visual interpretation, which mainly include two aspects: the detection range and the existence of saliency features. Inspired by the above concepts, superpixel segmentation is integrated into RN estimation. As presented in Fig. 1, the proposed approach consists of two steps: 1) estimation of local RN distribution and 2) rerectification of the registered image. For the first step, the superpixels are segmented from the input preregistered VHR image and sparse superpixels are generated by filtering all the superpixels in the frequency domain. Spatial correlation contrast is employed for the sparsely represented superpixels to estimate the RN distribution. For the second step, superpixel RN information is utilized to optimize the control points obtained in the preregistration process, and then the local transformation model is built by the previously optimized control points. The final fine-registered VHR image is generated by local rectification.

A. Generation of Sparse Superpixels

1) Superpixel Segmentation: Due to the influence of image distortion and complex ground objects, the matched control point error is randomly distributed in the range of the VHR image. Thus, the deviation between the registered image and the reference image shows different quantities. Superpixel algorithms group pixels into perceptually meaningful atomic regions by the degree of similarity among the neighboring pixels, which provides an effective basic object to calculate the local misalignments between two images. In the proposed method, simple linear iterative clustering [19] is used for superpixel segmentation, because it is computationally efficient and can generate superpixels compactly with uniform size and well adhered region boundaries.

The superpixel size is determined by the initial width S, which is assigned in advance during the segmentation process. Since the superpixels are used as the basic objects for RN estimation, while the local residual misalignments are randomly distributed in the registered image, the value of S has a very important impact on the accuracy and efficiency of the estimation process. When the S value is too small, the superpixel may contain only one kind of ground object, which can lead to a lack of saliency and can affect the rough location recognition. On the other hand, it is difficult to obtain precise position due to the large number of pixels and the complex changes of ground objects when the S value is too big. Further detailed analysis of the S value is presented in Section III.

2) Superpixel Sparse Representation: Like artificial recognition, it is easier to detect RN in high-frequency regions, which usually represent the edge regions and exhibit large differences in the spectral signature. Unfortunately, the low-frequency regions are not conducive to the calculation of global spatial correlation because of their similar spectral values. In the proposed method, the difference between original and prediction (DBOP) filter, which is a step of the Laplacian pyramid [20], is adopted to generate a bandpass frequency map of the preregistered image, while avoiding scrambling frequencies. The bandpass frequency map generated by DBOP filter is defined as

$$X_F = |X - M \uparrow (M \downarrow (X \otimes g_{5\times5})) \otimes g_{5\times5}| \tag{1}$$

where X represents the input registered image, $g_{5\times5}$ represents the 5×5 window of Gaussian filter kernel, and M is the sampling matrix, which is used to obtain the prediction image by downsampling and upsampling steps.

In order to obtain effective RN estimation objects, sparse representation of superpixels is carried out by frequency filtering according to the following two requirements: 1) redundant low-frequency pixels are removed and 2) each superpixel retains a sufficient number of pixels. Each superpixel specifies at least 50% of the pixels by calculating an adaptive frequency threshold, which is defined as follows:

$$T_F^n = \min\left\{ t \left| \sum_{i=1}^{t} P_{X_F}^n(i) > 0.5, t \in [0, X_{F\max}] \right. \right\} \tag{2}$$

where i is the value in X_F, n is the label of the superpixel, t is the current cumulative X_F value, and $P_{X_F}^n(i)$ denotes the frequency ratio of i calculated by the histogram statistic. Then, the threshold segmentation is employed to eliminate pixels that have a value of less than T_F^n in each superpixel. Letting (x, y) be the spatial position of the sample, the sparse superpixel is defined as follows:

$$\mathrm{SS}^n(x, y) = \begin{cases} X(x, y) & \text{if } X_F(x, y) \geq T_F^n \\ \emptyset & \text{else} \end{cases} \tag{3}$$

B. Estimation of RN Distribution Based on Sparse Superpixel

The estimation of RN distribution is actually the quantitative detection of local residual misalignments, which includes both intensity and direction misalignments. Since RN represents the drift of homonymous pixels between multitemporal images, a spatial correlation measure can be employed to estimate it.

We use a sparse superpixel here as the basic object and set a $w \times h$ search window for supporting the multiple displacement analysis in the reference image, where w and h are the width and height of the displacement range, respectively. In addition, the corresponding projected position on the reference image is obtained by bilinear interpolation resampling.

Supposing that the number of sparse superpixels in the registered image is N, for each sparse superpixel SS^n ($n = 1, \ldots, N$), the maximum correlation location of the reference image is searched. This can be regarded as the actual homonymous location and has two components $\{u^n, v^n\}$ within the search window coordinate system. However, it can be further improved to the subpixel level by coefficient balancing. The fine displacement coordinate is defined as follows:

$$u^{n'} = \begin{cases} u^n + \left(\dfrac{C^n}{C_l^n + C^n} - 1\right) \times \Delta d & \text{if } C_l^n > C_r^n \\ u^n + \dfrac{C^n}{C_r^n + C^n} \times \Delta d & \text{if } C_l^n < C_r^n \end{cases} \tag{4}$$

$$v^{n'} = \begin{cases} v^n + \left(\dfrac{C^n}{C_t^n + C^n} - 1\right) \times \Delta d & \text{if } C_t^n > C_b^n \\ v^n + \dfrac{C^n}{C_b^n + C^n} \times \Delta d & \text{if } C_t^n < C_b^n \end{cases} \tag{5}$$

where Δd is the moving step distance of each displacement, C^n is the maximum correlation coefficient, and C_l^n, C_r^n, C_t^n, and C_b^n denote the correlation coefficients of the left, right, top, and bottom sides of the location of C^n, respectively. Accordingly, the intensity ρ^n and direction θ^n of the local RN can be computed as

$$\rho^n = \sqrt{(u^{n'})^2 + (v^{n'})^2} \tag{6}$$

$$\theta^n = \tan^{-1}\left(\frac{u^{n'}}{v^{n'}}\right). \tag{7}$$

By applying the above operations to all the sparse superpixels, we can obtain an RN map of the preregistered image, which is useful for the subsequent fine registration.

C. Optimization of Initial Control Points

As mentioned in the previous registration process, the control points are obtained by the image matching algorithm introduced in [12], and their gross errors are calculated by RANSAC-based mismatch detection. For each RN estimated superpixel, the vector of misalignment is calculated. This information can be used to adjust the coordinates of the control points that are within this superpixel range. For each control point pair, refinement is applied to the transformed control point coordinates on the preregistered image, while the control point coordinates on the reference image are maintained. By combining the initial control points and the superpixel RNs, the detailed optimization procedure is shown as follows.

1) Selection of Estimated Noise Superpixels: Since the optimization is aimed at the poorly registered area and the RN vector is difficult to be accurately estimated in the regions with low correlation, not all the superpixels are available for optimization. RN intensity threshold T_ρ and maximum correlation coefficient threshold T_c are set to classify the superpixels into the nonselected and selected groups with the following formula:

$$\mathrm{SS}_U^n = \begin{cases} 0 & \text{if } \rho^n < T_\rho \;\&\&\; C^n < T_c \\ 1 & \text{else} \end{cases} \tag{8}$$

2) Selection of Effective Control Points: For the control points that fall within the selected superpixels, some are matched correctly, and those whose gross errors are greater than the index threshold are selected as the effective control points for further optimization.

3) Adjustment of Control Points: For the selected control points on the preregistered image, their new coordinates are obtained according to the estimated RN information of the superpixel in which they are located. These coordinates are considered as new control points that still correspond to the maintained control points on the reference image.

This approach obtains precisely corresponded control point pairs, and thus the local displacement relationship between the registered image and reference image is established.

D. Local Rectification

Since nonlinear local distortions exist in VHR images, it is unreasonable to apply transformation model (e.g., affine transformation) directly on the whole image for rectification. In order to mitigate the local residual misalignments in the preregistered image, local affine transformation based on triangulated irregular network is employed. For each of the triangles, the triangulated corresponding control points in the two images are used to calculate a warping function. Then, affine transformation is applied to the local rectification from the registered image to the reference image. Finally, the fine-registered image is generated after all the triangles are rectified.

III. Experiments and Analysis

A. Data Sets Description and Experimental Settings

To assess the effectiveness of the proposed approach, two multitemporal VHR remote sensing image data sets were used for experiments. The first data set (DS1), acquired over the region of Wenlin Town, China, contained two temporally different panchromatic images taken from GF-2 with a spatial resolution of 0.8 m. The reference image of DS1 was taken in April 2015, whereas the input image was taken in March 2016. A subset of 6000×6000 pixels that contains buildings, roads, farmland, and water area [Fig. 2(a) and (b)] was considered in the experiment. The second data set (DS2) is a set of multisource panchromatic images that are located at Shenzhen, China. The reference image of DS2 was taken from ZY-3 with a spatial resolution of 2.1 m in August 2015, whereas the input image was taken from GF-1 with a spatial resolution of 2 m in January 2017. Similarly, the above used subsets were both 6000×6000 pixels and contain mountains, vegetation, urban areas, and water areas [Fig. 2(c) and (d)].

Before applying the proposed approach, the input images were standardly registered by the ISS method [12] as previously mentioned. In addition, the influence of the S value setting was analyzed to select an appropriate segmentation size for SRNE. To this end, 30 pairs of randomly distributed points were manually selected in each of the data sets. Then, the deviations between the registered image and the reference image are distinguished by visual interpretation. The behavior of the RN estimation error by varying the S parameters is shown in Fig. 3. As can be seen, these two different resolution

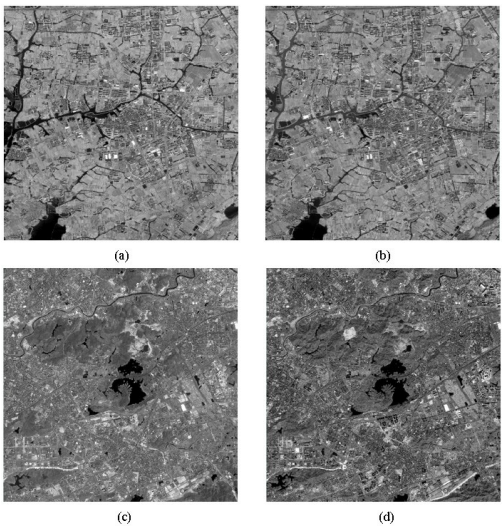

Fig. 2. Overview of two data sets. (a) and (b) GF-2 reference image and the GF-2 input image of DS1, respectively. (c) and (d) ZY-3 reference image and the GF-1 input image of DS2, respectively.

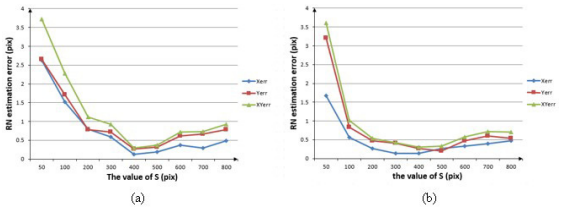

Fig. 3. Effects of the superpixel width S in (a) DS1 and (b) DS2. Xerr, Yerr, and XYerr represent the RN estimation error of X-direction component, Y-direction, and the total amount, respectively.

data sets both exhibit a smaller estimation error when the value of S was between 300 and 700, which is in line with the content we proposed in Section II-A. In our experiment, the superpixel width was set to be 400, which is applicable to different VHR images because this number of pixels is more likely to have enough significant features to support the spatial correlation analysis.

To quantitatively evaluate the registration performance, three indexes were used: 1) the root-mean-square error, which was calculated over 30 checkpoints that are manually extracted by experienced image interpreters from each data set and considered 0.5 pixels of artificial error; 2) the mean square of noise intensity (MSNI), which was calculated over the estimated superpixel RN by the formula $(\sum_{n=1}^{N} \rho^n)/N$; and 3) the unqualified alarm (UA) rate, which indicates the number of superpixels that are not conforming to the precision requirement over the total number of superpixels. By using these indexes, our proposed method was compared with three state-of-the-art methods: M UR-SIFT (feature-based standard registration) [6], ISS (preregistration process) [12], and SBFR (ISS preregistration) [16]. All the parameter settings of each method follow the authors' suggestions. In addition, an analysis was conducted on the impact of performing or nonperforming the SSR step when the proposed approach is employed. The parameters recommended for our approach were set as follows. The setting of moving step distance Δd and RN intensity threshold T_ρ influences the accuracy of RN estimation. As the VHR image registration required a precision of 1 m in most applications, Δd and T_ρ were determined by the resolution R of the input image, where $\Delta d = T_\rho = 1/R$.

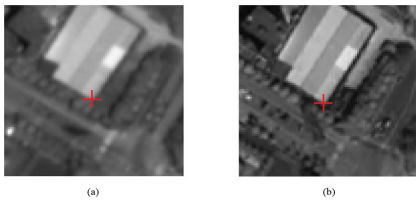

Fig. 4. Example of checkpoint on the (a) reference image and its projection point on the (b) SRNE fine registered image.

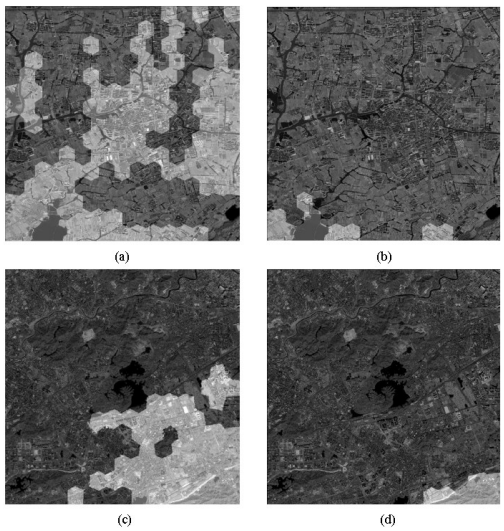

Fig. 5. RN distribution superposition diagram of two data sets. (a) and (b) RN distributions of the GF-2 image (a) before and (b) after applying the proposed approach. (c) and (d) RN distributions of the GF-1 image (c) before and (d) afterapplying the proposed approach.

The maximum correlation coefficient threshold T_c was set to be 0.3, which can effectively remove invalid estimation areas.

B. Results and Analysis

By using ISS matching results and the proposed control points optimization approach, 1521 and 1715 corresponding control point pairs were obtained in DS1 and DS2, respectively. After completing the proposed fine registration process, SRNE was performed on the fine-registered image once again. Fig. 4 shows an example of checkpoint on the reference image and its projection point on the SRNE fine-registered image. The RN distribution superposition diagram generated before and after applying the proposed fine registration on the two data sets is shown in Fig. 5. The estimated noise superpixels correspond to the light regions, whereas the non-noise superpixels correspond to the dark regions. It can be seen that the sizes of the noise regions after applying the proposed approach [Fig. 5(b) and (d)] were significantly less than the preregistration results [Fig. 5(a) and (c)]. As expected, the overall registration accuracy in the two data sets was effectively improved. For further visual comparison, Fig. 6 displays four groups of detailed checkerboard images generated from the two experimental data sets. The red circles in Fig. 6 highlight some of the edge objects that make it easy to evaluate the registration results. The misalignments generated from the preregistration process [Fig. 6(a)–(d)] were effectively eliminated in the fine-registered images generated by the proposed method [Fig. 6(e)–(h)], where the edge objects were precisely aligned with each other.

Table I shows the quantitative assessments of different registration methods for the two data sets. As one can see,

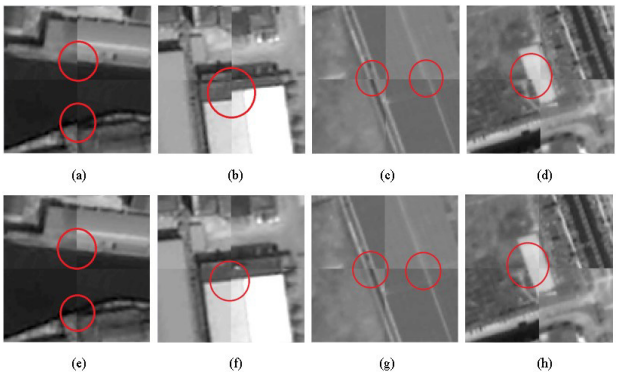

Fig. 6. Detailed checkerboard images generated from two data sets. (a) and (b) From DS1 and (c) and (d) from DS2 are the preregistered local subscene, whereas (e)–(h) corresponding subscene after applying the proposed fine registration approach.

TABLE I

SUMMARY OF THE QUANTITATIVE ASSESSMENTS FOR DIFFERENT METHODS ON DS1 AND DS2

Dataset	Method	RMSE (pixel)	MSNI (pixel)	UA (%)
DS1	M UR-SIFT	3.13	3.68	50.65
	ISS	2.97	3.46	48.44
	SBFR	1.25	1.78	11.47
	SRNE without SSR	1.71	1.97	14.50
	SRNE	**0.88**	**0.85**	**4.89**
DS2	M UR-SIFT	1.35	1.46	10.29
	ISS	1.59	1.81	13.76
	SBFR	0.84	0.59	5.96
	SRNE without SSR	0.89	0.71	7.47
	SRNE	**0.73**	**0.21**	**2.75**

both fine registration approaches of SBFR and SRNE achieved excellent performance for the improvements compared with standard registration approaches of M UR-SIFT and ISS. In DS1, the ISS preregistration resulted in an MSNI value of 3.46 pixels and an UA value of 48.44%. After applying SBFR and the proposed approach, the MSNI values were improved to 1.78 and 0.85 pixels and the UA values were improved to 11.47% and 4.89%, respectively. In DS2, the overall registration accuracy was better than that of DS1 because it has lower ground resolution. The ISS preregistration solved the problem of control point generation from multisource imagery by achieving a MSNI value of 1.81 pixels and an UA value of 13.76%. SBFR and the proposed approach improved the MSNI to 0.59 and 0.21 pixels and improve UA to 5.96% and 2.75%, respectively. In addition, we can see that the registration becomes less accurate without performing the SSR step in both the two data sets. In general, the proposed fine registration approach effectively mitigates the misalignments in the preregistration results and achieves superior performance to the other methods for both data sets.

IV. CONCLUSION

This letter proposed a novel fine registration approach for VHR images capable of estimating and mitigating local residual misalignments. The proposed method employs SSR to robustly preserve the edge pixels and improve the RN estimation, which is employed to optimize the initial control points. The fine-registered VHR image is generated by local rectification transformation, which is calculated from the newly optimized control points. As a fine registration approach, the proposed method can effectively improve the accuracy of the standard registration result. Compared with SBFR, the proposed approach uses the combination of ISS matching results and RN information to obtain sufficient and accurate control points, which can be applied to most areas of the images. Both the qualitative comparison and quantitative assessment verify the effectiveness and practicability of the proposed approach. As far as its limitations, it is difficult to completely avoid RN estimation error, and different misalignments remain for some of the superpixels. Thus, how to generate more robust estimation objects is worthy of further study.

REFERENCES

[1] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vision Comput.*, vol. 21, no. 11, pp. 977–1000, 2003.
[2] J. P. Lewis, "Fast template matching," *Vis. Interface*, vol. 32, no. 4, pp. 351–361, 1995.
[3] 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.
[4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[5] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, vol. 3, 1988, pp. 147–151.
[6] S. Paul and U. C. Pati, "Remote sensing optical image registration using modified uniform robust SIFT," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 9, pp. 1300–1304, Sep. 2016.
[7] Q. Li, G. Wang, J. Liu, and S. Chen, "Robust scale-invariant feature matching for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 287–291, Apr. 2009.
[8] I. Misra, S. M. Moorthi, D. Dhar, and R. Ramakrishnan, "An automatic satellite image registration technique based on harris corner detection and random sample consensus (RANSAC) outlier rejection model," in *Proc. Int. Conf. Recent Adv. Inf. Technol.*, Mar. 2012, pp. 68–73.
[9] J. Ma, J. C. W. Chan, and F. Canters, "Fully automatic subpixel image registration of multiangle CHRIS/Proba data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2829–2839, Jul. 2010.
[10] H. Goncalves, L. Corte-Real, and J. A. Goncalves, "Automatic image registration through image segmentation and SIFT," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2589–2600, Jul. 2011.
[11] H. Gonçalves, J. A. Gonçalves, and L. Corte-Real, "HAIRIS: A method for automatic image registration through histogram-based image segmentation," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 776–789, Mar. 2011.
[12] X. Ling, Y. Zhang, J. Xiong, X. Huang, and Z. Chen, "An image matching algorithm integrating global SRTM and image segmentation for multi-source satellite imagery," *Remote Sens.*, vol. 8, no. 8, p. 672, 2016.
[13] L. Bruzzone and R. Cossu, "An adaptive approach to reducing registration noise effects in unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2455–2465, Nov. 2003.
[14] F. Bovolo, L. Bruzzone, and S. Marchesi, "Analysis and adaptive estimation of the registration noise distribution in multitemporal VHR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2658–2671, Aug. 2009.
[15] Y. Han, F. Bovolo, and L. Bruzzone, "Edge-based registration-noise estimation in VHR multitemporal and multisensor images," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 9, pp. 1231–1235, Sep. 2016.
[16] Y. Han, F. Bovolo, and L. Bruzzone, "Segmentation-based fine registration of very high resolution multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2884–2897, May 2017.
[17] Q. Chen, S. Wang, B. Wang, and M. Sun, "Automatic registration method for fusion of ZY-1-02C satellite images," *Remote Sens.*, vol. 6, no. 1, pp. 157–179, 2013.
[18] L. Yu, D. Zhang, and E.-J. Holden, "A fast and fully automatic registration approach based on point features for multi-source remote-sensing images," *Comput. Geosci.*, vol. 34, no. 7, pp. 838–848, 2008.
[19] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
[20] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.