Motion Consistency-Based Correspondence Growing for Remote Sensing Image Matching

Yizhang Liu, Yanping Li, Luanyuan Dai, Taotao Lai, Changcai Yang, Lifang Wei, and Riqing Chen

***Abstract*— In this letter, we propose a remote sensing image matching method that is simple yet efficient to deal with different deformations. Inspired by the region growing strategy used in image segmentation, we integrate the motion consistency into the general region growing pipeline from a novel perspective. Specifically, we first obtain a subset with a high ratio inlier as the seed correspondence set. Then, to find more reliable correspondences, we formulate the motion consistency into the correspondence growing criterion, which is general to be suitable to many remote sensing applications. Extensive experimental results on the public available remote sensing data set show that our method achieves the best performance compared with state-of-the-art methods.**

***Index Terms*— Image matching, motion consistency, region growing, remote sensing.**

I. Introduction

IMAGE matching is the prerequisite of many vision-based tasks (e.g., image mosaic [1], image fusion [2], and 3-D reconstruction [3]) in remote sensing scenery, which aims to align the given image pairs captured from different viewpoints, at different times or with different modalities. One of the popular strategies to deal with image matching is feature-based pipeline, and it mainly involves the following two steps [4]:

1) constructing initial putative matches based on the existing feature matching techniques, such as SIFT [5], SURF [6], and ORB [7];
2) seeking reliable correspondences from the putative sets by imposing some geometric constraints [8].

In this letter, we mainly focus on step 2). During the past several decades, a number of mismatch removal methods have been developed. RANSAC [9] is one of the most well-known resampling methods, which adopts a hypothesize-and-verify procedure to sample the minimal subset with all inliers to estimate parameters of the model by resampling. ICF [10] is a representative nonparametric interpolation method, and it learns a pair of correspondence functions based on some geometrical prior information. LFGC [11] is a deep learning-based method that is the first one to introduce deep learning into the mismatch removal problem. GMS [12] and LPM [13]–[15] are motion consistency-based methods that are based on the intuitive observation that correct matches have similar motion behavior.

Nevertheless, there still exists some challenges to be solved. Specifically, for resampling methods, they need a predefined model assumption, and it will be less efficient to deal with complex nonrigid transformation. For the nonparametric interpolation methods, they commonly have cubic complexities, which limits their applications on real-time tasks. Besides, the performance of the above two techniques is significantly affected by the inlier ratio of the putative set. With a decreased inlier ratio, the performance of these methods will be largely degraded, and vice versa. Although deep learning can extract more abstract information from the given data, its performance substantially depends on the specific data that are used to train. Therefore, it is not robust enough to deal with different image types and transformations. It is worth noticing that the formulation of motion consistency-based methods is quite simple; therefore, it can accomplish the mismatch removal from thousands of putative matches in a few milliseconds, and it can also be suitable for various image transformations. However, sometimes, the current motion consistency-based methods cannot achieve satisfactory matching results because of filtering out some inliers by mistake. Therefore, to solve this problem, a more robust and better precision–recall tradeoff method is desirable.

In this letter, we propose a remote sensing image matching method from a novel perspective. Inspired by the region growing strategy used in image segmentation, we integrate the motion consistency into the general region growing pipeline. The schematic illustration of our proposed method is shown in Fig. 1. We first construct initial putative matches, as shown in Fig. 1(a), which contains both inliers (marked in yellow) and

Manuscript received October 12, 2020; revised December 8, 2020; accepted December 21, 2020. Date of publication January 20, 2021; date of current version January 4, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61701117, Grant 61702101, Grant 61802064, and Grant 61501120; in part by the Natural Science Fund of Fujian Province under Grant 2019J01402; in part by the Fujian Province Health Education Joint Research Project under Grant 2019WJ28; and in part by the Science and Technology Innovation Special Fund Project of Fujian Agriculture and Forestry University under Grant CXZX2019124S and Grant CXZX2019125G. *(Corresponding author: Changcai Yang.)*

Yizhang Liu is with the College of Computer and Information Science, Digital Fujian Research Institute of Big Data for Agriculture and Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China, and also with the School of Software Engineering, Tongji University, Shanghai 201804, China (e-mail: lyz8023lyp@gmail.com).

Yanping Li is with the Department of Computer Science and Technology, Tongji University, Shanghai 201804, China (e-mail: liyp8023yz@163.com).

Luanyuan Dai, Changcai Yang, Lifang Wei, and Riqing Chen are with the College of Computer and Information Science, Digital Fujian Research Institute of Big Data for Agriculture and Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China (e-mail: dailuanyuan@qq.com; changcaiyang@fafu.edu.cn; weilifang1981@163.com; riqing.chen@fafu.edu.cn).

Taotao Lai is with the College of Computer and Control Engineering, Minjiang University, Fuzhou 350108, China (e-mail: laitaotao@gmail.com).

Digital Object Identifier 10.1109/LGRS.2020.3048258

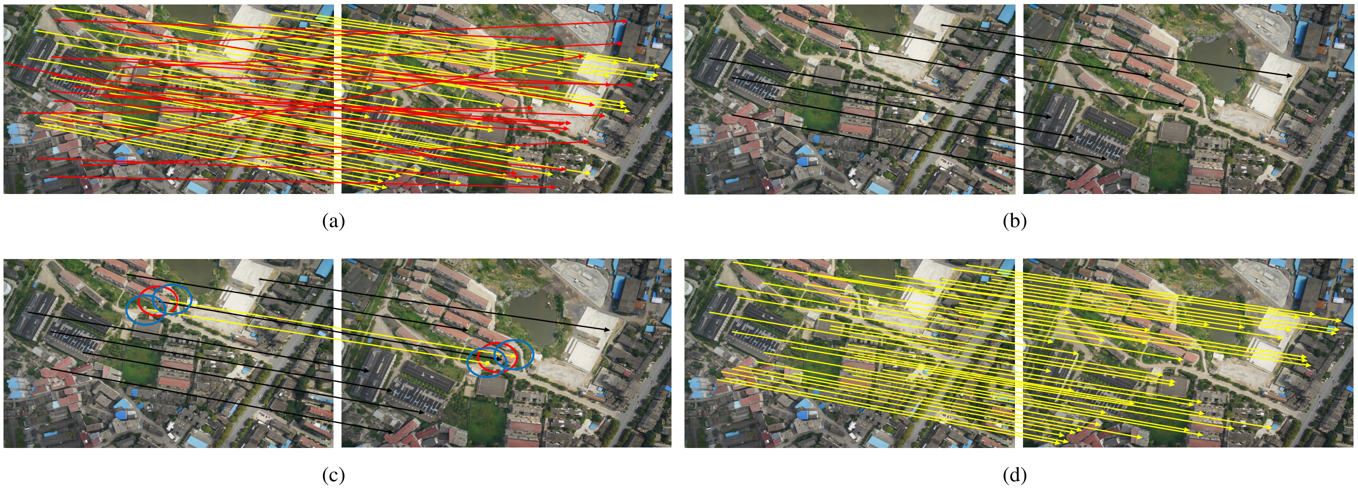


Fig. 1. Schematic illustration of our method on two remote sensing images. (a) Initial putative matches, where yellow and red arrows indicate inliers and outliers, respectively. (b) Initial seed correspondences (marked in black). (c) Process of correspondence growing, where red and blue circles indicate the neighborhood of initial seed correspondences and the newly added seed correspondences, respectively. (d) Final results obtained by the process of correspondence growing terminated.

outliers (marked in red). Then, we obtain a subset with high ratio inliers as the initial seed correspondences by imposing strict motion consistency constraint on the constructed putative sets, as shown in Fig. 1(b). After that, we formulate the motion consistency into a correspondence growing criterion. Specifically, if a putative match satisfies the criterion, it will be regarded as a new seed correspondence, as shown in Fig. 1(c). Finally, when the growing process of all seed correspondences terminates, all the seed correspondences constitute the inlier set, as shown in Fig. 1(d).

Overall, the contributions of this letter are twofold: 1) we propose a correspondence growing framework for remote sensing image matching and integrate the motion consistency into this framework and 2) we redefine two important issues of region growing technique in the content of feature matching and give a general solution.

II. Method

In this letter, we first obtain M putative matches $S = \{(x_i, y_i)\}_{i=1}^{M}$ extracted from two remote sensing images by using SIFT [5], where x_i and y_i are the spatial coordinates of the corresponding feature points. Besides, we use K nearest neighbors under the Euclidean distance to represent the region around a feature point. There are two important issues in the correspondence growing method, namely, the initial seed correspondences, and the growing criterion and process. In the following, we will introduce them in detail.

A. Selection of Initial Seed Correspondences

As we know, a suitable selection of initial seed points has an important impact on the final image segmentation results. Likewise, in the proposed framework, a good initialization will benefit the final matching results. In the context of feature matching, seed points are replaced by the seed correspondences; therefore, we aim to find a subset of some reliable correspondences from the initial putative set to initialize our algorithm. During the past several decades, there are many well-designed techniques that can achieve this goal [3], [5], [16], [17]. Nevertheless, many repetitive structures in the remote sensing images make the performance of these descriptor information-based methods largely degraded. Specifically, the ratio test strategy of the nearest neighbor and the second nearest neighbor will filter out most inliers between the similar structures (with similar descriptors), which will not be beneficial to the process of correspondence growing. We show an example in Fig. 2. x_1 is a feature point in the left image, and y_1 is its corresponding feature point in the right image according to the nearest neighbor strategy. Due to similar neighborhood information, these five feature points (x_1, y_1, y_2, y_3, and y_4) have similar descriptors. According to the ratio test strategy [5], the correct match (x_1, y_1) will be filtered out and not be regarded as a seed correspondence leading to the failure of finding other potential correct matches (marked in blue) [18]. Therefore, we propose a method of selecting initial seed correspondences by using motion consistency. Generally, it is clear that an inlier will have more supporters (other inliers) in its neighborhood due to similar motion behavior. Conversely, an outlier will have no or fewer supporters (other outliers) since the mistakes have randomly appeared, as shown in Fig. 1(a). Hence, the motion consistency for a putative match (x_i, y_i) can be formulated as

$$\Re_i = \mathbb{N}_i / K \tag{1}$$

Fig. 2. Illustration of the defect of ratio test strategy in the case of repetitive structures.

Fig. 3. Example of outliers in the initial seed correspondences.

where $\mathbb{N}_i$ is the number of neighboring matches with respect to the putative match (x_i, y_i) and K is the number of nearest neighbors of feature points. We denote the neighborhood of feature point x_i and y_i as N_{x_i} and N_{y_i}, respectively. It is easy to know that $\Re_i \in [0, 1]$. Ideally, for an inlier (x_i, y_i), $\Re_i$ is close to 1, and for an outlier (x_j, y_j), $\Re_j$ is close to 0. Therefore, we can use a predefined threshold λ to obtain a subset with high ratio inliers as seed correspondence set

$$S_\iota = \{x_i | \Re_i > \lambda, i = 1, 2, \ldots, M\}. \tag{2}$$

B. Growing Criterion and Process

To find more correct matches, the process of correspondence growing starts from the initial seed correspondences to adjacent correspondences with similar motion properties according to a similarity criterion. In fact, two nearby inliers have little difference both in length and angle under the Euclidean distance due to physical constraint, and it will be quite different for outliers [19]. We quantize the difference by designing a weighted distance $d(v_i, v_j)$ as follows:

$$d(v_i, v_j) = \Upsilon(v_i, v_j) + \xi \cdot \Theta(v_i, v_j) \tag{3}$$

with

$$\Upsilon(v_i, v_j) = \frac{\max\{|v_i|, |v_j|\}}{\min\{|v_i|, |v_j|\}} - 1 \tag{4}$$

where v_i is the replacement vector of putative match (x_i, y_i) (an initial seed correspondence) and v_j is the replacement vector of its neighboring match (x_j, y_j). $\Upsilon(v_i, v_j)$ is the ratio of length of v_i and v_j. $\Theta(v_i, v_j)$ is the angle between v_i and v_j. ξ is a weighted coefficient to trade off these two terms. Ideally, if v_i and v_j have high motion consistency, $d(v_i, v_j)$ will be close to 0, and if they have no or less motion consistency, $d(v_i, v_j)$ will be relatively far away from 0. Therefore, we can use a predefined threshold τ to determine whether a putative match (x_j, y_j) can be regarded as a seed correspondence according to the following rules:

$$c(v_j) = \begin{cases} 0, & d(v_i, v_j) \geq \tau \\ 1, & d(v_i, v_j) < \tau. \end{cases} \tag{5}$$

Specifically, if $c(v_j) = 1$, the putative match (x_j, y_j) will be newly added as a seed correspondence; otherwise, it will not be regarded as a seed correspondence in current growing process. In fact, the seed correspondences may contain outliers due to relatively loose motion consistency constraint. We show an example in Fig. 3. In this case, both $\mathbb{N}_i$ and K are equal to 3, and $\Re_i$ is equal to 1. It is easy to know that the putative match (x_i, y_i) will be selected as a seed correspondence according to (2). It is obvious that (x_i, y_i) is an outlier (marked in red), and the mistake is caused by ignoring the neighborhood topology consensus of seed correspondence. Therefore, we denote ℓ_i as the number of newly added seed correspondences associated with the putative match (x_i, y_i), and to enhance the robustness, we determine whether a seed correspondence is inlier by comparing ℓ_i with a predefined threshold α as follows:

$$p(v_i) = \begin{cases} 0, & \ell_i < \alpha \\ 1, & \ell_i \geq \alpha. \end{cases} \tag{6}$$

Specifically, $p(v_i) = 1$ points to an inlier, and $p(v_i) = 0$ denotes an outlier. Hence, the final inlier set $\mathcal{I}^*$ can be obtained with no seed correspondence increased, which can be determined by

$$\mathcal{I}^* = \{v_i | \ p(v_i) = 1, i = 1, \ldots, M\}. \tag{7}$$

Algorithm 1: Pseudocode of the Proposed Method

Input: putative set $S = \{(x_i, y_i)\}_{i=1}^{M}$,parameters K,λ,ξ,τ,α
Output: inlier set $\mathcal{I}^*$
Initialize $j = 0$;
Construct neighborhoods $\{N_{x_i}, N_{y_i}\}_{i=1}^{M}$ based on S;
Obtain a subset with high ratio inlier S_j using Eq. (2);
repeat
  Construct neighborhoods $\{N_{x_i}, N_{y_i}\}_{i=1}^{M}$ based on S_j;
  $j = j + 1$;
  Obtain a subset with high ratio inlier S_j using Eq. (2);
until $j \geq 2$;
Determine an initial seed correspondence set $S_\iota = S_j$;
Correspondence growing using Eq. (5);
Calculate $\{p(v_i)\}_{i=1}^{M}$ using Eq. (6);
Determine the inlier set $\mathcal{I}^*$ using Eq. (7);

C. Completed Proposed Method

The proposed method is summarized in Algorithm 1, which, from a novel perspective, formulates the feature-based remote sensing image matching as a motion consistency-based correspondence growing method. In order to obtain a more reliable seed correspondence set, we adopt a coarse-to-fine strategy to obtain a subset with high ratio inliers in case of heavy outliers. Specifically, we first construct the neighborhood for each feature point based on the initial putative set S. Then, we filter out some distinct outliers using (2) with $\lambda = 0.1$ obtaining a subset and use it for the next round of neighborhood construction of each feature point. It is clear that a large value of λ will increase the inlier ratio and simultaneously decrease the number of inliers of the subset and vice versa. Therefore, we repeat the above process with $\lambda = 0.3$ and $\lambda = 0.5$ in the next two rounds, respectively. It is clear that the large value of K, λ, and α will increase the precision and simultaneously decrease the recall of the initial seed

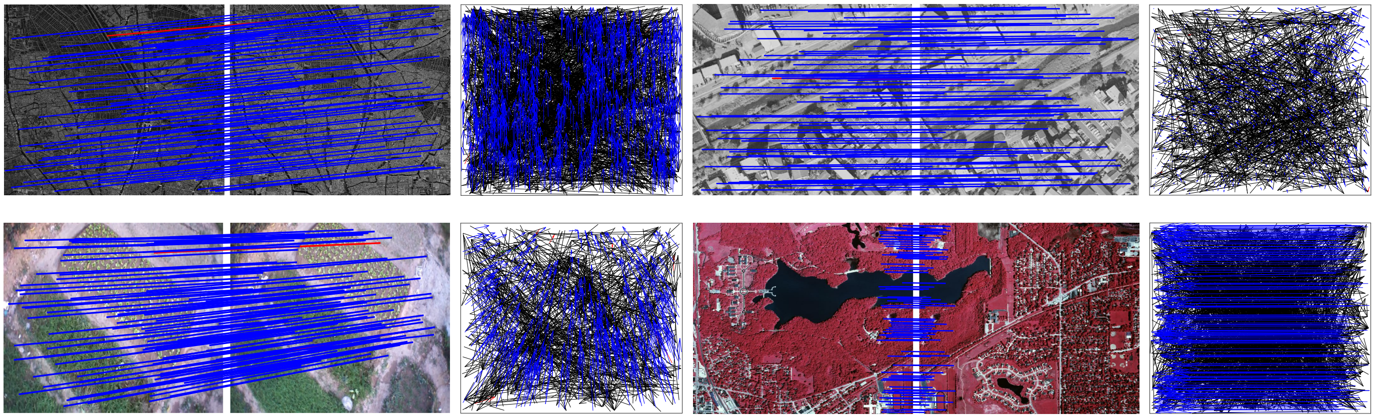
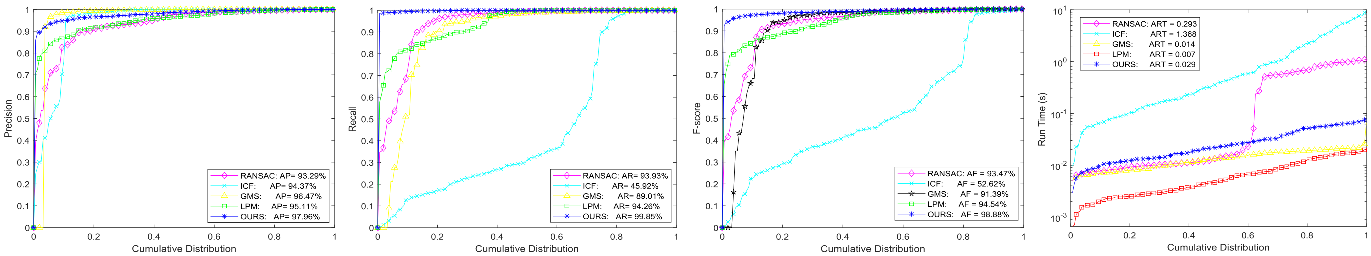


Fig. 4. Feature matching results of our method in four typical remote sensing image pairs. For each image pair, we show the correspondences identified by our method and its corresponding motion field (blue = true positive; black = true negative; red = false positive; and green = false negative). For visibility, no more than 100 matches are randomly selected to be shown.



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

Fig. 5. Quantitative comparisons of RANSAC [9], ICF [10], GMS [12], and LPM [13] with our proposed method. A point with coordinates (x, y) on a curve of (a), (b), (c) and (d) indicates there are $(100 * x)\%$ of image pairs, and their precision, recall, F-score and run time do not exceed y, respectively.

correspondences. With obtained seed correspondence set S_t, we can achieve the process of correspondence growing by using (5) and determine the inlier set using (6) and (7). In our evaluation, we set the default values as $K = 20, 10$, and 9 in the three rounds, respectively, and $\xi = 0.1$, $\tau = 0.15$, and $\alpha = 3$.

III. Experimental Results

In this section, we test the performance of our proposed method on the public available remote sensing data set [20], in which there are 161 image pairs involving four image types: synthetic aperture radar (SAR) images, panchromatic aerial photographs (PAN), unmanned aerial vehicle (UAV) images, and color infrared aerial photographs (CIAP). The challenges include severe noise, ground relief variations, multiple similar structures, and nonrigid transformation. The average inlier ratio and match number with respect to the data set are 68.76% and 918.32, respectively. All experiments are conducted on a laptop with Windows 10 operating system, Intel 1.60 GHz, 16-GB RAM, and MATLAB code. We use precision, recall, F-score, and run time to describe the performance and efficiency of the methods.

We first present the intuitive matching results on four typical image pairs (SAR049, PAN083, UAV007, and CIAP126) in Fig. 4 and the detailed statistical data in Table I. These image pairs suffer from affine distortion, small overlap, projective distortion, and rigid transformation, respectively. From Fig. 4 and Table I, we can see that the proposed method obtains satisfactory matching results especially in recall statistics, which benefits from the process of correspondence growth. Our proposed method can even solve the feature matching problem in the case of an extremely low inlier ratio in the putative set (e.g., image pair CIAP126), which benefits from the strategy of selection of the initial seed correspondences.

TABLE I

Statistical Data (%) With Respect to the Typical Image Pairs of Our Proposed Method

Image pair	*Inlier ratio*	*Precision*	*Recall*	*F-score*
SAR049	39.90	98.78	99.22	99.00
PAN083	57.57	98.72	100.0	99.36
UAV007	42.18	98.45	99.74	99.09
CIAP126	12.03	100.0	100.0	100.0

Then, we show the quantitative comparisons of RANSAC [9], ICF [10], GMS [12], and LPM [13] with our proposed method on the whole data set in Fig. 5. Carefully observing Fig. 5, we can find our proposed method achieve higher precision statistics than RANSAC, LPM, and LMR in most cases but only slightly lower than ICF in a few cases [see the upper left corner of Fig. 5(a)]. Note that, for the recall statistics [see Fig. 5(b)], we achieve the best results than any other competitor, which also demonstrates the effectiveness of the motion consistency-based correspondence growing strategy. Overall, our method obtains the best precision–recall tradeoff reflected in the F-score statistics [see Fig. 5(c)]. In addition, from Fig. 5(d), we can find that our proposed method can accomplish mismatch removal within tens of milliseconds over 1000 putative matches and is much faster than RANSAC and ICF. Since our proposed method focuses on finding more accurate inliers by several iterations, the computational complexity is relatively larger than GMS and LPM.

IV. Conclusion

In this letter, we propose a motion consistency-based correspondence growing method for remote sensing image matching, which is designed from a novel perspective by combining the simple yet effective motion consistency with the region growing framework. We redefine two important issues of region growth in the context of feature matching and present a general solution. Specifically, a subset with a high ratio inlier as the seed correspondence set gives the better initialization and makes the algorithm quickly converged. The growing criterion is based on intuitive observation and does not rely on any predefined transformation, and hence, it can be suitable for many remote sensing tasks. The qualitative and quantitative results demonstrate the superiority of our proposed method over state-of-the-art competitors.

In the future, we will investigate another strategy to obtain a more reliable seed correspondence set and effective correspondence growing criterion. Besides, we will combine our proposed algorithm with other high-level applications, such as remote sensing image fusion and target detection.

References

[1] F. Fang, T. Wang, Y. Fang, and G. Zhang, "Fast color blending for seamless image stitching," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 7, pp. 1115–1119, Jul. 2019.

[2] K. Sun, L. Liu, and W. Tao, "Image matching via feature fusion and coherent constraint," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 3, pp. 289–293, Mar. 2017.

[3] W.-Y. Lin *et al.*, "CODE: Coherence based decision boundaries for feature correspondence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 1, pp. 34–47, Jan. 2018.

[4] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," *Int. J. Comput. Vis.*, pp. 1–57, Aug. 2020.

[5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[6] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, May 2006, pp. 404–417.

[7] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.

[8] J. Ma, X. Jiang, J. Jiang, J. Zhao, and X. Guo, "LMR: Learning a two-class classifier for mismatch removal," *IEEE Trans. Image Process.*, vol. 28, no. 8, pp. 4045–4059, Aug. 2019.

[9] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[10] X. Li and Z. Hu, "Rejecting mismatches by correspondence function," *Int. J. Comput. Vis.*, vol. 89, no. 1, pp. 1–17, Aug. 2010.

[11] K. M. Yi, E. Trulls, Y. Ono, V. Lepetit, M. Salzmann, and P. Fua, "Learning to find good correspondences," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2666–2674.

[12] J.-W. Bian *et al.*, "GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence," *Int. J. Comput. Vis.*, vol. 128, no. 6, pp. 1580–1593, Jun. 2020.

[13] J. Ma, J. Zhao, J. Jiang, H. Zhou, and X. Guo, "Locality preserving matching," *Int. J. Comput. Vis.*, vol. 127, no. 5, pp. 512–531, May 2019.

[14] X. Jiang, J. Ma, J. Jiang, and X. Guo, "Robust feature matching using spatial clustering with heavy outliers," *IEEE Trans. Image Process.*, vol. 29, pp. 736–746, 2020.

[15] Y. Liu *et al.*, "Robust feature matching via advanced neighborhood topology consensus," *Neurocomputing*, vol. 421, pp. 273–284, Jan. 2021.

[16] J. Ma, J. Jiang, H. Zhou, J. Zhao, and X. Guo, "Guided locality preserving feature matching for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4435–4447, Aug. 2018.

[17] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.

[18] C. Yang, M. Zhang, Z. Zhang, L. Wei, R. Chen, and H. Zhou, "Non-rigid point set registration via global and local constraints," *Multimedia Tools Appl.*, vol. 77, no. 24, pp. 31607–31625, 2018.

[19] A. L. Yuille and N. M. Grzywacz, "A mathematical analysis of the motion coherence theory," *Int. J. Comput. Vis.*, vol. 3, no. 2, pp. 155–175, Jun. 1989.

[20] X. Jiang, J. Jiang, A. Fan, Z. Wang, and J. Ma, "Multiscale locality and rank preservation for robust feature matching of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6462–6472, Sep. 2019.