



Local neighbor propagation on graphs for mismatch removal

Hanlin Guo ^{a,b,c}, Guobao Xiao ^{d,*}, Lumei Su ^a, Jiaxing Zhou ^a, Da-Han Wang ^e

^a School of Electrical Engineering and Automation, Xiamen University of Technology, China

^b Xiamen Key Laboratory of Frontier Electric Power Equipment and Intelligent Control, China

^c Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, China

^d Key Laboratory of Complex Systems Modeling and Simulation, School of Computer Science and Technology, Hangzhou Dianzi University, China

^e Fujian Key Laboratory of Pattern Recognition and Image Understanding, School of Computer and Information Engineering, Xiamen University of Technology, China

ARTICLE INFO

Keywords:

Feature matching
Geometric model fitting
Outlier removal
Local neighbor propagation
Clustering

ABSTRACT

Recently, despite the promising progress having been achieved, most mismatch removal methods only consider the connection relationships of feature matches in a single neighborhood, and ignore their relationships between different neighborhoods, which will lead to the unavailability of local topological structure for feature matching. In this paper, we propose a novel robust Local Neighbor Propagation on Graphs based mismatch removal (LNPG) method for robust feature matching. LNPG starts from a novel neighborhood graph construction strategy, which leverages both the spatial and the residual information to preserve the local neighborhood structures of potential inliers. Subsequently, LNPG incorporates local neighbor propagation into the graph to enhance connection relationships of the data in different neighborhoods, by using the path-based similarity measurement and the adaptive graph partition. In addition, LNPG includes a novel consistency-filtering-based clustering algorithm, which introduces a reliable neighborhood consistency measure function and an effective cluster merging criterion for robust clustering. Overall, LNPG not only effectively distinguishes inliers from outliers, but also reliably classifies inliers into different transformation models between pairs of images. Extensive experiments on publicly available datasets show the superiority of our LNPG in comparison with other state-of-the-art methods.

1. Introduction

Feature matching aims to establish reliable feature correspondences between two images of the same scene captured from different perspectives, at different times, or by different sensors. It is a prerequisite for many computer vision and pattern recognition applications, including image retrieval, image registration, and motion segmentation, remote sensing [1–4]. In general, the problem of feature matching is solved in a two-stage manner: constructing putative feature matches between given two images and removing false matches from them. In the first stage, putative feature matches are usually generated by simply picking out pairs of feature points (such as, Harris [5]) with similar feature descriptors (such as, SIFT [6]). Nevertheless, the putative feature matches often contain a large number of outliers (i.e., false matches also called mismatches) in addition to inliers (i.e., true matches), due to the

* Corresponding author.

E-mail address: gbx@mju.edu.cn (G. Xiao).

imperfection of feature point detector and ambiguities of local descriptors. Thus, in the second stage, it is critical to remove outliers while preserving inliers in the putative feature matches as much as possible.

A variety of mismatch removal methods have been developed during the past few decades [7–10], and most of them use the principle of preserving structures of potential inliers to effectively remove outliers. While these methods are effective for general cases of mismatch removal, their performance dramatically degenerates for the data with high outlier ratios. The reason behind this is that, on the one hand, the neighborhood sizes of data are hard to determine for the complex cases (such as, imaging viewpoint changes and multimodel data). On the other hand, outliers are also easy to be selected as neighbors with inappropriate neighborhood sizes. As a result, it is hard to obtain the convincing effect in neighborhood structure preservation for mismatch removal. For example, previous mismatch removal methods [11,12] rely on the connection relationships of feature matches in each single neighborhood for mismatch removal, while they are poor in the connections of feature matches between different neighborhoods, leading to suboptimal results of mismatch removal.

To overcome the above challenges, we propose a novel robust Local Neighbor Propagation on Graphs based mismatch removal method (called LNPG), to accurately remove outliers from putative feature matches between two-view images. We observe that, while image pairs suffer from rotation, scale variance, rigid or non-rigid transformation, the spatial neighborhood structures of inliers usually tend to be preserved well; But the spatial neighborhood structures of outliers tend to be considerably different. Based on this observation, we present a novel neighborhood graph construction strategy for robust feature matching. Specifically, we first generate a series of neighborhood sets based on the spatial consistency constraint and then construct a neighborhood graph by using residual information, to preserve the local neighborhood structures of potential inliers. Here, residual information is derived from the residual values between input data and the generated model hypotheses from input data. Then, we propose a local neighbor propagation strategy to boost the connections of the data in different neighborhoods, by using the path-based similarity measurement and the entropy-threshold-based graph partition. This is because that the captured neighborhood sets of potential inliers inevitably contain the outliers as outliers randomly scatter across images, and this will produce unconvincing geometrical structures for matching results. Based on the local neighbor propagation strategy, LNPG is able to boost the topological connections of the data within each neighborhood and capture the global consistency among non-neighbors through the data in different neighborhoods. As a result, the influence of outliers can be effectively reduced and the reliability of the captured local topological structures can be improved for the mismatch removal task.

In addition, we propose a novel consistency-filtering-based clustering algorithm to effectively classify input data into inliers belonging to different transformation models and outliers. The clustering algorithm not only includes a reliable neighborhood consistency measure function for cluster detection, but also includes an effective cluster merging criterion for cluster merging.

The contributions of this paper are summarized as follows:

- We construct a novel neighborhood graph fully utilizing both spatial and residual information to preserve local neighborhood structures of potential inliers for outlier removal. This graph can help the proposed method effectively eliminate outliers while retaining inliers with high robustness.
- We propose a local neighbor propagation strategy that utilizes path-based similarity measurement and adaptive graph partitioning to enhance connections of the data in different neighborhoods and accurately represent local structures. Consequently, the local topological structures of potential inliers can be preserved effectively for mismatch removal.
- We propose a novel consistency-filtering-based clustering algorithm, to simultaneously label the input data as inliers or outliers and classify the inliers into different transformation models between pairs of images, without requiring the predefined cluster numbers.

The qualitative and quantitative experiments on publicly available datasets with different types of image transformations show that the proposed method is able to achieve the state-of-the-art performance for mismatch removal.

The remainder of this article is organized as follows: In Section 2, we introduce necessary background information and give an overview of the related work. In Section 3, we describe the proposed mismatch removal method in detail. In Section 4, we provide the experimental results on publicly available datasets for the feature matching task and the two-view geometry estimation task. Finally, we give the conclusions in Section 5.

2. Related work

In general, the problem of feature matching is typically solved in a two-stage manner, i.e., generating putative feature matches and removing false matches from them. In this section, we review the related methods for feature matching according to these two stages.

For the first stage, salient feature points and their descriptors are captured from each image, and the putative feature matches are constructed through measuring the local feature descriptors. There are many popular feature descriptors including SIFT [6], and ORB [13]. SIFT is one of the most popular descriptors, and it is widely used in the feature matching problem due to its invariance to rotation, scale, and illumination. Some improvements have been proposed based on SIFT to make it more appropriate for feature matching. For example, CoFSM [14] detects feature points in a new image scale space by considering the co-occurrence filter for generating sufficient putative feature matches. These methods can effectively improve the quality of the generated putative feature matches. However, the captured matches are inevitably contaminated with a large number of mismatches due to the ambiguities of local feature descriptors. Therefore, to effectively solve the feature matching problem, the second stage is designed to remove

mismatches from the captured putative matches. In this paper, we mainly focus on the mismatch removal problem. For a brief review, we roughly divide the mismatch removal methods into the three categories, i.e., learning-based methods, nonparametric-based methods, and parametric-based methods.

In recent years, deep learning techniques have been developing rapidly, and many learning-based mismatch removal methods have also been proposed for robust feature matching. LFGC [15] is a pioneering work to introduce a learning-based technique by following the PointNet-like architecture for mismatch removal. LFGC is able to achieve promising solutions with embedding global information in feature matches, but it discards the useful local spatial information. Then, NM-Net [8] uses a network architecture similar to LFGC and adds the compatibility-specific locality information to improve the network performance. MS2DG-Net [16] learns to find true matches according to the semantics information between the captured feature matches from given pairs of images. SuperFusion [17] incorporates image registration, fusion, and semantic segmentation into a single framework to robust feature fusion. PGFNet [18] adopts a preference-guided filtering strategy to learn correspondences. Although the learning-based methods are able to achieve high matching accuracy, the generalization capacity of these methods is still in need of study. This is because that massive labeled training data should be demanded for these methods to guarantee their performance. Moreover, this demand may severely restrict their applications in many practice scenarios.

Nonparametric-based methods typically exploit specified functions based on various local information as constraints to distinguish true matches from false matches. For example, GMS [9] exploits motion smoothness information in a statistical manner to remove mismatches. LPM [12] exploits the local neighborhood information of feature matches to find consistent matching. The above-mentioned methods are able to obtain high computational speed in the solutions. However, these methods merely use the connection relationships in each single neighborhood and the connections between different neighborhoods are ignored, which is insufficient to depict local topological structures.

Parametric-based methods usually use global information as constraints (such as, affine transformation model [19]) for mismatch removal. They commonly perform the model hypothesis generation and model verification steps to seek a representative model hypothesis as the estimated model instance. The feature matches that are consensus with the estimated model instance are selected as true matches. RANSAC [20] is one of the most well-known methods, which alternately performs the above two steps to maximize the size of the consensus set (i.e., inlier set). Due to the appealing performance of RANSAC, some improved versions (e.g., [21,22]) have been proposed for robust feature matching. The authors in [23] propose a consensus sampling technique to increase the probability of sampling inliers for robust and efficient feature matching. EAS [24] uses robust loss functions to perform deterministic search for the inliers or geometric model. These methods are effective for general cases of mismatch removal, however, their performance is unreliable for the data highly contaminated with outliers. In addition, to handle the data with multiple structures, many variants of RANSAC have been proposed to improve the matching performance. They classify the data into outliers and the inliers corresponding to different structures, such as, RCMSA [25], RansaCov++ [26], MSHF [27], RFMSCAN [28], IFTLN [29], and CBG [30]. However, only a few methods fully take advantage of local neighborhood information to cluster data with multiple structures for robust feature matching problems.

The proposed method (LNPG) in this paper is one of the parametric-based methods, and it fully takes advantages of local neighborhood information and the overlapping information between different neighborhoods for mismatch removal. LNPG utilizes the local neighbor propagation to enhance the connections between different neighborhoods and the topological connections within neighborhoods, and thus the effectiveness of outlier removal can be improved significantly. Moreover, LNPG can provide reliable matching results by proposing a novel consistency-filtering-based clustering algorithm to effectively label the data into inliers or outliers. In addition, we distinguish inliers from outliers according to the data clustering results, by which the sensitivity to the inlier threshold for the classification of feature matches can be significantly alleviated.

3. Method

In this section, we describe the details of the proposed LNPG method for mismatch removal. To this end, we start by constructing a set of putative feature matches established by feature extraction and feature description (such as, Harris [5] and SIFT [6], respectively). Specifically, we preserve the local neighborhood structures of potential inliers via a robust neighborhood graph representation in Section 3.1. Then, we enhance the local connections of the data between neighborhoods, through local neighbor propagation along paths of the graph derived from the residual information in Section 3.2. Finally, we propose a novel clustering algorithm by cluster detection and cluster merging to classify the putative matches into false matches and true matches in Section 3.3.

3.1. Local neighbor preservation via neighborhood graph representation

Suppose that we have obtained the input data $S = \{s_n\}_{n=1}^N$ of N feature matches, where $s_n = (x_n, y_n)$ is a feature match (i.e., a datum), and x_n and y_n are the spatial positions of two feature points, respectively. Our goal is to accurately distinguish inliers from outliers in the input data for robust matching.

In real-world data, the local neighborhood relationship of a true match will get well preserved even in the presence of large viewpoint changes, because their corresponding feature points in one image often lie in the same object [10,25]. Therefore, true matches often have many similar neighboring matches, whose corresponding feature points in pairs of images are respectively close to each other, while false matches not. Based on the observation, we propose to reliably compute the neighborhood sets of feature matches and effectively determine the neighborhood sizes by searching similar neighboring matches.

Given an image pair, let a feature match $s_n = (x_n, y_n) \in S$ be a reference datum, x_n and y_n be two feature points that connect to the feature match s_n . Formally, we define the similar neighboring matches of s_n as follows:

$$\mathcal{N}_{s_n} = \{s_{n'} = (x_{n'}, y_{n'}) | x_{n'} \in \mathcal{N}_{x_n}, y_{n'} \in \mathcal{N}_{y_n}\}, \quad (1)$$

where \mathcal{N}_{x_n} and \mathcal{N}_{y_n} respectively denote the K -nearest neighbors of the two feature points x_n and y_n , and \mathcal{N}_{s_n} is regard as the neighborhood set of s_n , and $0 \leq |\mathcal{N}_{s_n}| \leq K$. As can be seen, the selection of neighborhood sizes of feature matches is driven by data based on the locally spatial relationships of feature points. To effectively characterize the complex relationships among input data with outliers, we construct a neighborhood graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_n\}_{n=1}^N$ denotes the set of vertices corresponding to the data in $S = \{s_n\}_{n=1}^N$, and \mathcal{E} denotes the set of edges between pairs of connected vertices. In \mathcal{G} , every vertex is locally connected by an edge to each of its similar neighboring matches according to Eq. (1). After that, the neighborhood graph \mathcal{G} is associated with a weighted adjacency matrix $\mathbf{W} = [w_{(v_n, v_{n'})}]$. Here, $w_{(v_n, v_{n'})}$ is the similarity between v_n and $v_{n'}$, and it is computed based on residual information between two vertices in the following.

Given the input data $S = \{s_i\}_{i=1}^N$, we sample a set of minimal subsets from S and generate corresponding M model hypotheses $\mathcal{H} = \{\mathbf{h}_m\}_{m=1}^M$, for representing the relationships between data and model hypotheses. For each datum s_n , we compute its residual $r(s_n, \mathbf{h}_m)$ with respect to m -th model hypothesis \mathbf{h}_m based on the Sampson distance [31]. Then, let us introduce the preference function [32], which indicates the degree of the preference of a datum s_n to a model hypothesis \mathbf{h}_m , as follows:

$$f_m^{(n)} = \begin{cases} e^{-r^2(s_n, \mathbf{h}_m)/\delta^2}, & \text{if } r(s_n, \mathbf{h}_m) < \tau_m, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

where τ_m is an inlier scale estimated by IKOSE [33], and δ is a normalization constant as in [34]. As a result, a preference vector $\mathbf{f}^{(n)}$ for a datum s_n towards the set of M model hypotheses \mathcal{H} is written as $\mathbf{f}^{(n)} = [f_1^{(n)}, f_2^{(n)}, \dots, f_M^{(n)}]$. After that, the residual correlation associated with s_n and $s_{n'}$ is computed based on the preference vectors $\mathbf{f}^{(n)}$ and $\mathbf{f}^{(n')}$ by using the cosine similarity as:

$$\varphi_{(s_n, s_{n'})} = \frac{\langle \mathbf{f}^{(n)}, \mathbf{f}^{(n')} \rangle}{\|\mathbf{f}^{(n)}\| \times \|\mathbf{f}^{(n')}\|}, \quad (3)$$

where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the inner product and the corresponding induced norm, respectively. As can be seen in Eq. (3), if the two data (s_n and $s_{n'}$) are the inliers, $\varphi_{(s_n, s_{n'})}$ is high because these two data share many common model hypotheses. Otherwise, $\varphi_{(s_n, s_{n'})}$ is low because the insignificant model hypotheses sampled from outliers start to dominate.

Based on the residual and spatial information, the similarity $w_{(v_n, v_{n'})}$ between the vertices v_n and $v_{n'}$ is computed as:

$$w_{(v_n, v_{n'})} = \begin{cases} \varphi_{(s_n, s_{n'})}, & s_{n'} \in \mathcal{N}_{s_n}, \\ 0, & s_{n'} \notin \mathcal{N}_{s_n}, \end{cases} \quad (4)$$

where $\varphi_{(s_n, s_{n'})}$ denotes the residual correlation between s_n and $s_{n'}$ using Eq. (3). Unlike the full-connected graph representing each vertex by all vertices, the constructed neighborhood graph in the proposed LNPG representing each vertex by only its similar neighboring vertices. Thus, LNPG can effectively preserve the local neighborhood structures of potential inliers, and alleviate the influence of outliers, which are usually outside the local structures of potential inliers.

3.2. Local neighbor propagation via paths in graphs

The constructed neighborhoods by the spatial consistency constraint would lead to unsatisfactory matching results when handling complex data heavily contaminated with outliers. Because the connection relationships of the data in single neighborhood are insufficient to yield the availability of local topological structures. Reliable matching results are root in the sufficient within-neighborhood and between-neighborhood information for local and global structures. Therefore, a neighborhood propagation strategy is introduced into the proposed matching methods based on the paths in graphs, to boost the topological connections within neighborhoods and connections between different neighborhoods. Specifically, we first introduce the path-based similarity measurement to consider the similarities between any two vertices including non-adjacent vertices in a global sense, to robustly estimate the edge weighting scores in a graph. Then, we present the entropy-threshold-based graph partition by using the information theory to remain the graph edges with high weighting scores while remove the other ones. One of the most significant advantages is that neighbor propagation could enhance topological connections between inliers while alleviating the influence of outliers. As a result, we are able to obtain reliable neighborhoods during the graph partition process. Fig. 1 shows an example of local neighbor propagation.

3.2.1. Path-based similarity measurement

Based on the captured neighborhoods, the proposed matching methods introduce the local neighbor propagation via path-based similarity measurement into graphs to improve matching performance. Paths characterize the connectivity of the network associated with a graph. Some studies (e.g., [35–38]) have shown that the paths in a graph are favorable for expressing complex relationships among vertices. In this paper, we define a path-based similarity measurement on the neighborhood graph \mathcal{G} associated with the weighted adjacency matrix $\mathbf{W} = [w_{(v_n, v_{n'})}]$. More specifically, we firstly denote a path of length t that passes through the vertices $(v_1, v_2, v_3, \dots, v_t, v_{t+1})$ on the graph \mathcal{G} between vertices v_1 and v_{t+1} as $\gamma_t = \{v_1 \rightarrow v_2 \rightarrow v_3 \dots v_t \rightarrow v_{t+1}\}$. Here, the vertices v_1 and v_{t+1}

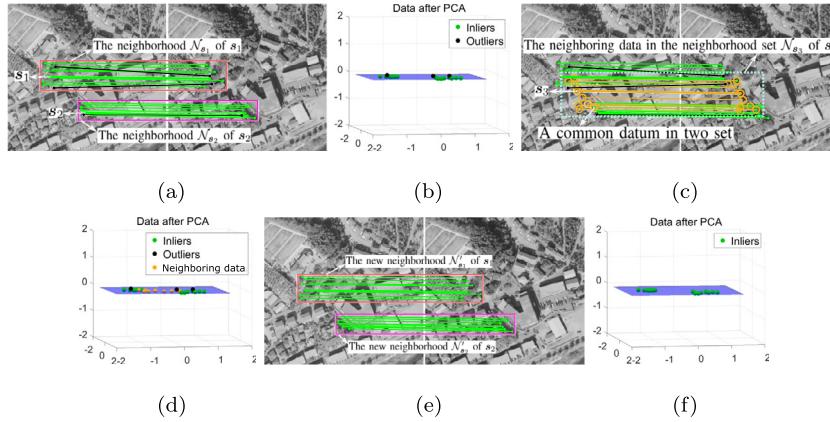


Fig. 1. An example of local neighbor propagation among the data for outlier removal. (a), (c) and (e) show the feature matches in the image pairs. Note that, the boldfaced lines denote the reference datum of a neighborhood set. The green and yellow colored lines denote inliers, while the black colored lines denote outliers. (b), (d) and (f) exhibit the 3D visualization by using PCA. (a) The data in N_{s_1} is apart from the data in the N_{s_2} . (c) The data in N_{s_1} and N_{s_2} have no neighborhood relations, while their similarities can be effectively measured based on neighbor propagation exploiting the path similarities through neighboring data in N_{s_3} . (e) The outliers in N_{s_1} and N_{s_2} can be removed by neighbor propagation and graph partition, enhancing the topological connections within each neighborhood.

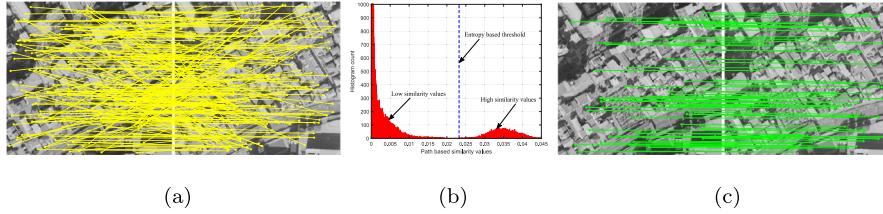


Fig. 2. Distribution of the estimated similarity values obtained by the path-based similarity measurement. (a) Input data. (b) Distribution of the obtained similarity values. (c) The estimated inliers by the proposed mismatch removal method. From (b), we can see that the estimated similarity values appear as two sharp peaks with high separability. An entropy based threshold obtained by information theory can be used to separate the similarities with low similarity values and high similarity values.

corresponds to the data s_n and $s_{n'}$, respectively. Then, the path-based similarity on a specific path γ_t is defined as $p_{\gamma_t} = \prod_{\ell=1}^t w_{(v_\ell, v_{\ell+1})}$. However, it is possible that there are more than one path of length t between vertices v_n and $v_{n'}$. Thus, denoting all possible paths of length t between v_n and $v_{n'}$ as P_t , the t -path similarity $p_t(v_n, v_{n'})$ can be defined as follows:

$$p_t(v_n, v_{n'}) = \sum_{\gamma_t \in P_t} p_{\gamma_t}(v_n, v_{n'}). \quad (5)$$

The t -path similarity $p_t(v_n, v_{n'})$ is used to measure the similarity associated with the two vertices v_n and $v_{n'}$ at t -path scale. However, the paths at different path scales (i.e., $t = \{1, 2, \dots, \infty\}$) should be integrated to effectively measure similarities between any two vertices. To integrate all possible paths, we make use of the generating function from [36,38] and define the generating function with regularization, which can be used to estimate a significant similarity value. The recently proposed optimization methods [39–41] can help effectively calculate the generating function. The generating function for the t -path similarities is defined as follows:

$$q(v_n, v_{n'}) = \sum_{t=1}^{\infty} \alpha^t p_t(v_n, v_{n'}), \quad (6)$$

where α is a regularization parameter as in [38], and α^t can be regarded as the weight for t -path similarity. After that, the value of the generating function $q(v_n, v_{n'})$ can be regarded as the (n, n') -th element of the propagation weight matrix $\mathbf{Q} \in \mathbb{R}^{N \times N}$ as [38]:

$$\mathbf{Q} = (\mathbf{I} - \alpha \mathbf{W})^{-1} - \mathbf{I}, \quad (7)$$

where \mathbf{I} represents the identity matrix, and \mathbf{W} denotes the weighted adjacency matrix. Here, the element $q(v_n, v_{n'})$ in the matrix \mathbf{Q} denotes the path-based similarity value (i.e., edge weighting score) between two vertices v_n and $v_{n'}$. Based on the constructed neighborhood graph \mathcal{G} , LNPG propagates neighborhoods and defines propagation weight matrix \mathbf{Q} by computing the path-based similarities between two vertices including non-adjacent vertices.

3.2.2. Entropy-threshold-based graph partition

The similarity values in \mathbf{Q} are high if they are between two vertices corresponding to the inliers; The similarity values in \mathbf{Q} are low, otherwise (see Fig. 2 for an example with two sharp peaks). This is because that the residual correlations between the inliers

tend to be obviously larger than those between outliers and outliers or inliers and outliers according to Eq. (3). Therefore, we can implement a simple yet effective graph partition step by using a threshold value, to remain the edges with high weighting scores (i.e., high similarity values between two vertices), while remove the other ones. As a result, the global structural information of the graph (i.e., the edge connections between the inliers) can be preserved.

However, it is hard to manually set a proper threshold value for different data during the graph partition process. In this paper, we propose to use the information theory algorithm [42] to adaptively capture a threshold for the graph partition. By the information theory, we perform the binary classification operation on the elements from the matrix \mathbf{Q} to generate a binary adjacency matrix $\hat{\mathbf{Q}}$, which is used to reform the connection relationships between vertices in the neighborhood graph. More specifically, given the data $S = \{s_n\}_{n=1}^N$ and the matrix $\mathbf{Q} = [q(v_n, v_{n'})]$, let $\phi_{n,n'}^{(r)} = \max(\mathbf{Q}) - (q(v_n, v_{n'}))^2$ denote the gap between the maximum element of $\mathbf{Q} \in \mathbb{R}^{N \times N}$ and the r -th element $q(v_n, v_{n'})$ of \mathbf{Q} . Then, we can obtain the prior probability of $\phi_{n,n'}^{(r)}$ as:

$$\rho(\phi_{n,n'}^{(r)}) = \phi_{n,n'}^{(r)} / \sum_{g=1}^{N^2} \phi_{n,n'}^{(g)}. \quad (8)$$

The entropy of prior probability can be computed as:

$$E = - \sum_{r=1}^{N^2} \rho(\phi_{n,n'}^{(r)}) \log \rho(\phi_{n,n'}^{(r)}). \quad (9)$$

Here E is treated as a threshold to dichotomize all elements in \mathbf{Q} as follows:

$$\hat{q}(v_n, v_{n'}) = \begin{cases} 1, & \text{if } -\log \rho(\phi_{n,n'}^{(r)}) > E, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

An example of the captured entropy based threshold is shown in Fig. 2 (b).

After applying the above operations for all the elements, the matrix $\mathbf{Q} = [q(v_n, v_{n'})]$ is updated to a binary adjacent matrix $\hat{\mathbf{Q}} = [\hat{q}(v_n, v_{n'})]$, which signifies global relations between vertices in a graph. Then, we remain the edges between vertex v_n and its connected vertex $v_{n'}$ with $\hat{q}(v_n, v_{n'}) = 1$, and cut off the rest. By reconstructing the edges of the neighborhood graph \mathcal{G} , we effectively perform graph partition to obtain the partitioned graph \mathcal{G}' , and thus effectively preserve the local topological structures of potential inliers.

The initially constructed graph via local neighbor preservation aims at maintaining the topological connections within each neighborhood, while the partitioned graph via local neighbor propagation pays more attention to expand the neighborhood connections between different neighborhoods. Consequently, the connected vertex groups in the partitioned graph tend to correspond to the inlier clusters, while the isolated vertices are likely to correspond to outliers. Thus, the partitioned graph is more beneficial for removing outliers from input data.

3.3. Robust clustering by using local and global structure consistency

After obtaining the partitioned graph, we can proceed with data clustering for distinguishing inliers from outliers in input data. Some previous clustering algorithms (e.g., [38]) can be used to identify clusters over the graph constructed by neighborhood relationships in a greed manner, and thus classify the vertices connected by edges into the same cluster. However, there would be two situations that may happen in the data clustering process through the neighborhood graph: (1) The cluster of inliers may wrongly contain outliers; (2) The cluster corresponds to an inlier group maybe divided into different connection groups. Therefore, to effectively handle the aforementioned two challenging situations, we propose a novel consistency-filtering-based clustering (CFC) algorithm including cluster detection and cluster merging by making appropriate use of local and global relationships, to label the input data as inliers or outliers and classify the inliers into different transformation models.

3.3.1. Cluster detection via local structure consistency

In general, the two vertices have many common neighboring vertices if they correspond to inliers; They have few common neighboring vertices, otherwise. By exploring the common neighboring vertices on the partitioned graph \mathcal{G}' , we design a neighborhood consistency measure (*NCM*) function to estimate whether two neighboring vertices should be classified into the same cluster. Specifically, given the vertices $\mathcal{V} = \{v_n\}_{n=1}^N$ and their binary adjacency matrix $\hat{\mathbf{Q}} = [\hat{q}(v_n, v_{n'})] \in \mathbb{R}^{N \times N}$, the number of the neighboring vertices for each vertex can be computed by $\mathbf{b} = \hat{\mathbf{Q}}\mathbf{e}^\top$. Here \mathbf{e} denotes a vector with all elements as 1. On the other hand, the number of the common neighboring vertices between two vertices v_n and $v_{n'}$ is computed by $\hat{\mathbf{Z}} = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^\top$. Here, the diagonal elements in $\hat{\mathbf{Z}}$ are set to 0, because no vertex connects to itself. After that, we define the novel neighborhood consistency measure function $V(v_n, v_{n'})$ as follows:

$$V(v_n, v_{n'}) = \max \left\{ \frac{z_{n,n'}}{b_n}, \frac{z_{n,n'}}{b_{n'}} \right\}, \quad (11)$$

where $\{b_n, b_{n'}\} \in \mathbf{b}$ and it represents the number of the neighboring vertices connected to v_n or $v_{n'}$ by edges. $z_{n,n'} \in \hat{\mathbf{Z}}$ and it represents the number of the common neighboring vertices between v_n and $v_{n'}$. According to Eq. (11), we can effectively estimate the consistency between two connected vertices with edges, by measuring their common neighboring vertices. We show an example to illustrate the calculation process of the vertex neighborhood consistency on a neighborhood graph in Fig. 3. The calculated results from Fig. 3 (b) and (c) indicate that the $V(v_4, v_{19})$ between the vertices corresponding to inliers is higher than the $V(v_{15}, v_{17})$ between the

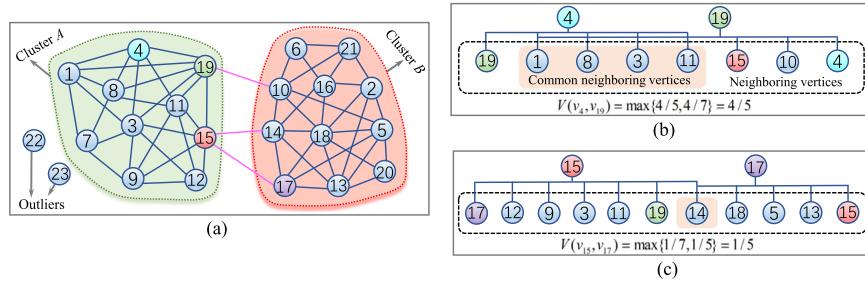


Fig. 3. An example of the proposed neighborhood consistency measurement for robust clustering. (a) The obtained neighborhood graph, where the vertices having the neighborhood relationships are connected by edges. (b) and (c) Two pair of vertices (i.e., 4 and 19, 15 and 17) with the neighboring vertices and their common neighboring vertices.

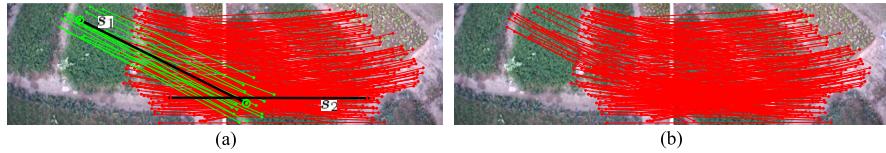


Fig. 4. The feature matching results obtained by the proposed method (only showing the obtained inliers by the proposed method). (a) The result before cluster merging. (b) The result after cluster merging. The inliers belonging to different subgroups are marked by the red, and green colors, respectively. The black color denotes the seeds in each subgroup.

vertices, which do not correspond to inliers. This shows the effectiveness of the proposed NCM for the neighborhood consistency measurement. After that, the defined NCM is used for the decision whether to assign the vertices into the same cluster during the clustering process.

Given the vertices $\mathcal{V} = \{v_n\}_{n=1}^N$, we denote the initial cluster index of each vertex as 0. For clustering, we first search two neighboring vertices over the partitioned graph \mathcal{G}' and look if NCM between the two vertices meets a threshold V . Then, by using the threshold V , we classify the vertices, whose NCM are higher than V into the same group, and denote them by the same cluster index. Finally, we repeat the above steps until all of the cluster indices associated with the neighboring vertices are updated. As a result, the cluster of the inliers is identified as the densely connected vertex group in the graph \mathcal{G}' . In contrast, most of the outliers are identified as the isolated vertices. However, in feature matching problems, not all the vertices corresponding to inliers are densely connected with each other. They may be united in a sparsely connected group of vertices, which may be divided into multiple densely connected groups and result in the over-segmentation problem (see Fig. 4 (a) for an example). Thus, it is necessary to merge the obtained subgroups that correspond to the inliers.

3.3.2. Cluster merging via global structure consistency

To handle the over-segmentation problem, we present a residual-based merging strategy by using the global residuals between the vertices (corresponding to the data) and the different subgroups (corresponding to the estimated transformation models). Based on the observation that the residual correlations between the inliers and an estimated transformation model tend to be high [25], we aim to use a residual-based measure to capture the intra-relationships of subgroups. We assume that there exists a seed in each subgroup, which can be used to reflect the residual distribution of the vertices in a subgroup (such as the data s_1 and s_2 in Fig. 4 (a)). Then, the variances of the residuals between the seeds towards two subgroups (before and after merging) can be used to decide whether the two subgroups should be merged into the same cluster or not. An example of subgroup merging obtained by the proposed method is shown in Fig. 4 (b).

Specifically, we select a seed from each subgroup according to the propagation weight matrix \mathbf{Q} in Eq. (7), which estimates the similarities between the vertices by all possible paths. The reason that we don't exploit the binary weighted matrix $\hat{\mathbf{Q}}$ is that based on the graph partition, the similarity in $\hat{\mathbf{Q}}$ is 0 for the vertices that are classified into different subgroups. Thus, it is unsuitable to decide whether two subgroups should be merged. With the weighted matrix \mathbf{Q} , the weighting score of each vertex is derived from the similarities between a vertex and all the other vertices, and it is computed as follows:

$$\omega(v_n) = \sum_{n'=1}^N q(v_n, v_{n'}) \quad (12)$$

After that, a seed is located by the datum with the highest weighting score in each subgroup. To compute residuals, we estimate a transformation model on the vertices from each subgroup. Let $r(v_\alpha, \mathcal{T}_\alpha)$ denote the residual between a seed v_α corresponding to a datum and a subgroup \mathcal{T}_α corresponding to an estimated model. We compute the residual $r(v_\alpha, \mathcal{T}_\alpha)$ based on the Sampson distance [31]. With the above definitions, we can proceed with the merging of subgroups. If the merging of two subgroups causes a small variance of the residuals associated with two seeds, the subgroups are supposed to correspond to the same transformation model. Thus, the merging criterion is computed as follows:

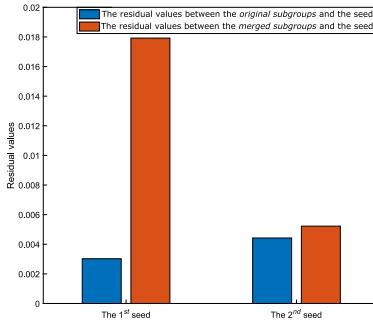


Fig. 5. The variances of the residuals between the seed towards two subgroups before and after merging.

Algorithm 1 Local Neighbor Propagation on Graphs for mismatch removal (LNPG) method.

Input: Feature matches $S = \{s_n\}_{n=1}^N$, parameters K and V .

Output: The labels of feature matches.

- 1: Generate a neighborhood set of each feature match from S based on the spatial consistency constraint by using Eq. (1).
 - 2: Construct a neighborhood graph \mathcal{G} based on residual and spatial information by using Eq. (4).
 - 3: Compute the path-based similarities of the graph \mathcal{G} by using Eq. (7).
 - 4: Partition the graph \mathcal{G} as a new graph \mathcal{C}' based on an adaptive threshold by using Eq. (10).
 - 5: Detect the clusters based on the novel neighborhood consistency measure function by using Eq. (11).
 - 6: Merge the clusters based on the designed merging criterion by using Eq. (13).
 - 7: Distinguish outliers from inliers belonging to different structures from feature matches according to the clustering result.
-

$$\zeta_{merge} = \left(|r(v_\alpha, \mathcal{T}_\alpha) - r(v_\alpha, (\mathcal{T}_\alpha \cup \mathcal{T}_\beta))| < 2\tau \right) \wedge \left(|r(v_\beta, \mathcal{T}_\beta) - r(v_\beta, (\mathcal{T}_\alpha \cup \mathcal{T}_\beta))| < 2\tau \right), \quad (13)$$

where $(\mathcal{T}_\alpha \cup \mathcal{T}_\beta)$ denotes two merged subgroups. 2τ is used to judge the variance of the residuals associated with seeds and subgroups, and τ is an inlier scale estimated by IKOSE [33]. This criterion checks the residual variances between two seeds and the subgroups before and after merging (as shown in Fig. 5), to judge if they produce small variances. The smaller residual variances mean that the estimated transformation model from the merged subgroup is close to the estimated transformation model from the original subgroups. Thus, if the criterion in Eq. (13) is satisfied, we merge two subgroups. By iteratively performing the subgroup merging between the subgroups, the proposed clustering algorithm (i.e., CFC) is able to accurately distinguish inliers from outliers. It is worth pointing out that CFC is also able to classify the input data into the inliers belonging to different transformation models and outliers for robust feature matching, when the input data involve multiple transformation models between pairs of images. In contrast to the entropy-threshold-based graph partition, CFC not only enables a more accurate segmentation of inliers and outliers, but also facilitates the assignment of inliers to different transformation models. CFC provides significant advantages in practical scenarios even in the presence of multiple structure data.

The time complexity of the proposed mismatch removal method (called LNPG) is mainly governed by searching the K -nearest neighbors of each feature point from input data and computing the matrix inversion, where the former requires $O((K + N)\log N)$ time complexity while the latter requires $O(N^{2.38})$ time complexity. Here, N is the number of feature matches from input data. In contrast, the other steps of LNPG take much less time than the above two steps. Therefore, the total time complexity of LNPG is about $O((K + N)\log N) + O(N^{2.38})$.

4. Experimental results

In this section, we test the performance of the proposed mismatch removal method (called LNPG) on different kinds of datasets, and we compare it with other representative state-of-the-art feature matching methods. We also use a publicly available dataset that includes single and multiple structure data for the performance evaluation. Specifically, we first introduce the datasets used in the experiments and the evaluation metrics to measure the matching performance. Then, we conduct experiments on various images to analyze the components of LNPG, including parameter analysis, and robustness test. After that, we evaluate the performance of the proposed LNPG by qualitative and quantitative experiments. All experiments are conducted on MS Windows with Intel Core i7-1260P CPU @2.1 GHz, 16 GB RAM.

4.1. Datasets and evaluation metrics

To conduct a comprehension evaluation for LNPG, we perform experiments on the four remote sensing datasets and a publicly available dataset as follows:

- 1) *CIAP*: The dataset contains 54 pairs of color infrared aerial photographs with the size 700×700 . The image pairs are taken from eastern Illinois, IL, USA. Note that, these image pairs are already orthorectified, and thus, the involved transformation model is

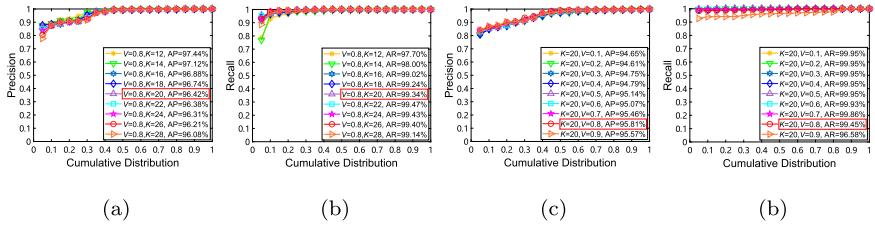


Fig. 6. Precision and recall with respect to the cumulative distribution obtained by the proposed method with different parameter settings on 20 remote sensing image pairs. (a) and (b) respectively denote precision and recall with the same $V = 0.8$ and different K . (c) and (d) respectively denote precision and recall with the same $K = 20$ and different V . The satisfactory average precision and recall are denoted by the boxes in the legends.

just rigid. However, the overlap areas between these image pairs are quite small. The image pairs from the CIAP dataset can be used to handle the image mosaic problem by the feature matching task.

2) *FE*: The dataset contains 18 pairs of color images with the sizes 1280×1024 or 1088×1088 . The image pairs are taken from different scenes by a fisheye camera. These image pairs suffer from viewpoint changes and heavy nonrigid deformations. The image pairs from the *FE* dataset can be used to handle the nonparametric image matching problem.

3) *PAN*: The dataset contains 33 pairs of panchromatic aerial photographs with the sizes 561×518 or 600×700 . The image pairs are taken at different time by a frame camera. These image pairs often involve ground relief variations or large viewpoint changes. The image pairs from the *PAN* dataset can be used to handle the change detection problem by the feature matching task.

4) *UAV*: The dataset contains 41 image pairs with the sizes 600×337 . The image pairs are taken from a piece of farmland by an unmanned gyroplane. These image pairs contain projective distortions. The image pairs from the *UAV* dataset can be used to handle the automatic crop monitoring problem by the feature matching task.

5) *AdelaideRMF*: The dataset contains 38 image pairs with the size 512×682 . In the dataset, the first 19 image pairs are taken from dynamic scenes with different moving objects, but the other 19 image pairs are taken from static scenes with different planes. The former can be used to handle the two-view based motion segmentation problem by the feature matching task, while the latter can be used to handle the two-view based plane segmentation problem by the feature matching task. It is worth pointing out that these image pairs from the *AdelaideRMF* dataset involve not only single structure data but also multiple structure data.

To evaluate the matching performance of competing methods, the commonly used metrics are precision, recall, and f-score. Precision is defined as the percentage of ground-truth true matches among all of the captured true matches identified by an algorithm. Recall is defined as the percentage of true matches identified by an algorithm among the whole ground-truth true matches. F-score is used to evaluate the comprehensive performance of matching considering both precision and recall. The precision, recall, and f-score are defined as follows:

$$\text{precision} = \frac{tp}{tp + fp}, \quad (14)$$

$$\text{recall} = \frac{tp}{tp + fn}, \quad (15)$$

$$f\text{-score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}, \quad (16)$$

where tp denotes the number of true positive matches; fp denotes the number of false positive matches; fn denotes the number of false negative matches.

In addition, the Clustering Error (*CE*) is used to measure the clustering performance when competing methods are used to deal with multiple structure data for robust feature matching. The defined *CE* is computed as follows:

$$CE = \frac{\text{the number of mislabeled data}}{\text{the total number of input data}}. \quad (17)$$

Note that, the lower value of *CE* denotes the better performance achieved by a feature matching method.

4.2. Analysis for the proposed LNPG

In this section, 20 image pairs with different kinds of image transformations including rotation, scale variance, rigid and non-rigid deformations are taken from the aforementioned datasets for evaluation, to validate the effectiveness of the proposed LNPG. The precision and recall are used to evaluate the performance of a mismatch removal method.

4.2.1. Parameter analysis and settings

There are two parameters used in the proposed LNPG, i.e., K determines the number of the nearest neighbors of each feature point, and it is used to generate similar neighboring matches. V is the cut-off threshold during the clustering process, and it is used to distinguish inliers from outliers. To seek the optimal values of these two parameters, i.e., K and V , we test different parameter settings on the 20 representative remote sensing image pairs. We report the precision and recall as shown in Fig. 6. As we can see, the performance usually presents a trend of increasing first and then decreasing for these two parameters with different values. To

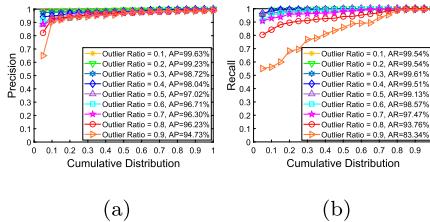


Fig. 7. Precision and recall with respect to the cumulative distribution obtained by the proposed method with different outlier ratios on 20 remote sensing image pairs.

Table 1

Average quantitative comparisons of LNPG and its four variants on 20 representative image pairs. AP-Average Precision; AR-Average Recall; AF-Average F-Score.

	LNPG	LNPG-V1	LNPG-V2	LNPG-V3	LNPG-V4
AP(%)	97.07	89.11	93.98	95.96	96.68
AR(%)	99.36	99.95	94.56	99.66	93.07
AF(%)	98.16	93.85	93.81	97.72	94.77

achieve a tradeoff between precision and recall, we empirically set the default values of these parameters as $K = 10$ and $V = 0.8$ in the subsequent evaluation for all the following experiments.

4.2.2. Robustness test

To test the robustness of the proposed mismatch removal method (LNPG), we evaluate the performance of LNPG on the 20 remote sensing image pairs with different ratios of outliers. We add the outliers into the used image pairs by the uniform distribution, and we vary the outlier ratios from 0.1 to 0.9 at an interval of 0.1. We change the number of outliers to each used image pair but fix the number of the initial inliers for experiments. We report the average precision and recall obtained by LNPG on these remote sensing images in Fig. 7. From the results, we can see that, the average recall does not show large fluctuations when the outlier ratios are between 0.1 and 0.8. The biggest gap of the average recall is 2.87%. However, the average precision shows no significant changes when the outlier ratios are increased from 0.1 to 0.9. Overall, the experimental results show the robustness of LNPG with regard to different outlier ratios.

4.2.3. Ablation study

Four variants are developed to analyze the influence of different components of the proposed LNPG. The four variants of LNPG are (1) LNPG-V1: LNPG without using the designed spatial consistency constraint to construct the neighborhood graph. It means that LNPG-V1 constructs the neighborhood graph by using the spatial relationship based on principal components analysis (PCA) as in RCMSA [25]; (2) LNPG-V2: LNPG without using the local neighbor propagation strategy; (3) LNPG-V3: LNPG without using the neighborhood consistency measure function; (4) LNPG-V4: LNPG without using the clustering merging criterion. In addition, LNPG, which uses all the investigated components, is regarded as a benchmark.

We evaluate the performance of LNPG and its four variants on the 20 representative image pairs and show the quantitative comparisons in Table 1. As can be seen, LNPG-V1 that does not use the spatial consistency constraint and LNPG-V2 that does not use the local neighbor propagation strategy show the worst performance. LNPG-V1 has a decrease in AF by 4.31%, while LNPG-V2 has a decrease in AF by 4.35%. Thus, both of them play significant roles in improving the robust performance of LNPG and the local neighbor propagation strategy contributes more to improve the performance. LNPG-V3 that does not use the neighborhood consistency measure (*NCM*) function has a decrease in AF by 0.44%, which demonstrates the effectiveness of the *NCM* for mismatch removal. Note that the AR of LNPG is lower than that of LNPG-V1 and LNPG-V3. The main reason is that LNPG is able to effectively distinguish inliers from outliers, while discarding a few inliers. In addition, LNPG-V4 that does not use the clustering merging criterion also has a relatively large performance deterioration in AF. This shows the effectiveness of the clustering merging criterion of LNPG for mismatch removal. Overall, LNPG also achieves the better performance in AP and AR than its four variants, which further shows the effectiveness of the investigated components of LNPG.

4.3. Results on feature matching

4.3.1. Qualitative results

We first give the qualitative matching results obtained by the proposed LNPG on some representative image pairs from the four datasets (i.e., CIAP, FE, PAN, UAV), as shown in Fig. 8. From top to bottom, each row shows two examples for a dataset. For each example, the intuitive matching result on an image pair is shown in the left plot, and the corresponding motion field of the matches is shown in the right plot. Note that, only 100 randomly selected matches of true positive, false negative and false positive are shown in the left plot for visibility, while the results of all matches in the motion field are shown in the right plot denoted by the arrows.

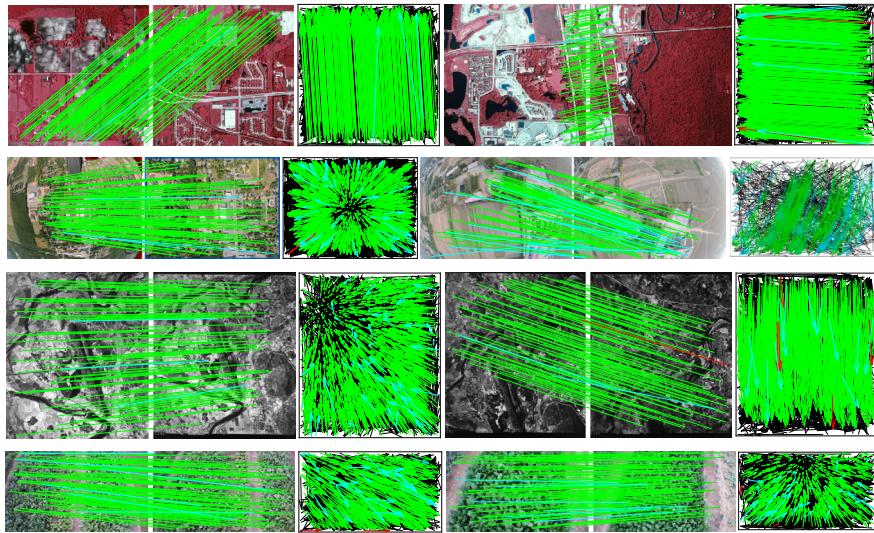


Fig. 8. Qualitative illustrations of feature matching results obtained by the proposed LNPG on eight representative remote sensing image pairs. From top to bottom and left to right: CIAP1, CIAP2, FE1, FE2, PAN1, PAN2, UAV1, and UAV2 with 14.23%, 19.30%, 31.97%, 46.90%, 60.42%, 39.05%, 31.97%, and 41.94% inlier ratios, respectively. For visibility, in each image pair, only 100 randomly selected matches are shown, but true negatives are not shown. We also show the motion field, in which the head and tail of each arrow respectively correspond to the positions of feature points in two images (green = true positive, black = true negative, red = false negative, cyan = false positive).

These representative image pairs involve many challenging situations, such as, small overlapping regions, nonrigid deformations, large viewpoint changes, and local geometric distortions, which brings a large number of outliers in initial feature matches from corresponding image pairs. For the input image pairs, we use the SIFT algorithm to extract features and generate feature matches that usually contains many outliers (i.e., false matches), except for inliers (i.e., true matches). For the eight image pairs, the number of inliers (and the inlier ratio) are 333(14.23%), 440(19.30%), 985(31.97%), 469(46.90%), 1334(60.42%), 786(39.05%), 392(31.97%), 546(41.94%), respectively. By using the proposed LNPG, we remove the false matches on the eight image pairs, and we obtain the precision, recall, and f-score as follows: (99.11%, 100.00%, 99.55%), (99.10%, 99.55%, 99.32%), (97.13%, 99.49%, 98.30%), (85.40%, 98.51%, 91.49%), (96.81%, 100.00%, 98.38%), (98.49%, 99.75%, 99.12%), (93.96%, 99.23%, 96.53%), (94.60%, 99.45%, 96.96%), respectively. As can be seen, LNPG is able to obtain good feature matching results on the eight challenging image pairs. LNPG successfully identifies most of the true matches from the input matches, and only a few matches are wrongly judged. These experimental results show the robustness and generality of the proposed method when handling the remote sensing image pairs with different types of transformations and a large number of outliers.

4.3.2. Quantitative results

To conduct quantitative comparisons, eight state-of-the-art methods for mismatch removal are adopted, including RANSAC [20], RANSAC++ [43], GS [44], GFC [11], LPM [12], mTopKRP [45], LAF [46], and LOGO [47]. We have attempted to optimize the parameters of all competing methods according to the suggestion of original papers for their best performance. We report the statistical results about precision, recall, F-Score, and runtime obtained by these competing methods on the four remote sensing datasets as shown in Fig. 9.

As can be seen, for the CIAP dataset, although mTopKRP and LAF achieve high accuracy, the proposed LNPG shows its ability to preserve more true matches, leading to better recall and f-score. RANSAC, LPM, and LOGO have high recall, but they have low precision and f-score. In contrast, GS and RANSAC++ have high precision and relatively low recall and f-score. GFC achieves worse matching results, and it has low f-score, due to its relaxed geometric constraint and the parameter sensitivity. For the FE dataset, it involves nonrigid deformations, and it is the most challenging dataset. All the nine competing methods cannot achieve better results for the FE dataset than those for the other three datasets. Nevertheless, the results show that, LNPG achieves the best recall among the three resampling-based methods (i.e., LNPG, RANSAC, RANSAC++) and it also achieves the best f-score among all the nine competing methods, mainly because of taking advantage of the local neighbor propagation strategy for mismatch removal. GS, GFC and LOGO are able to obtain better results than RANSAC and RANSAC++, while their performance is still worse than those of the proposed LNPG. Compared with mTopKRP and LAF, LNPG is able to preserve more true matches and obtain a better recall rate, due to the effectiveness of the local neighbor preservation and propagation. Although LPM would obtain comparative results, it can not work well on the data with low inlier ratios, resulting in the relatively low recall rate. For the PAN dataset, the proposed LNPG can obtain satisfying matching results on the most of image pairs with best values of recall and f-score and a good value of precision. The recently proposed mTopKRP, LAF and LOGO also achieve promising performance but their f-score is slightly worse than that of LNPG. In general, LNPG has better performance than RANSAC, GS, RANSAC++, GFC, and LPM because it takes into account both the connections of the data within neighborhoods and between different neighborhoods for mismatch removal. For the UAV dataset, we can observe that LNPG, mTopKRP, and LAF have the best performance. LPM and LOGO are less robust and fail in some of image

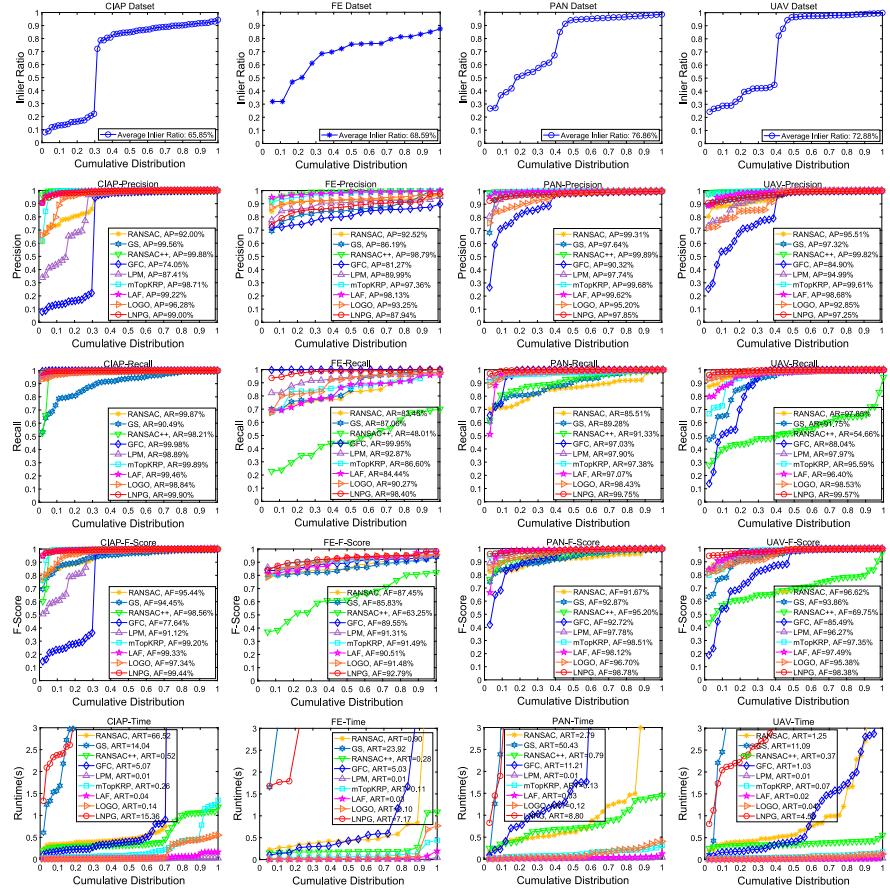


Fig. 9. Quantitative comparisons of RANSAC, GS, RANSAC++, GFCLPM, mTopkRP, LAF, LOGO and LNPG on four datasets that are CIAP, FE, PAN, UAV, and nonremote sensing, respectively (from top to bottom). Precision, recall, f-score and running time with respective to the cumulative distribution are shown from left to right.

pairs, resulting in relatively bad matching results. GS and RANSAC have relatively bad performance for this dataset that involves projective distortions and has serious outliers. RANSAC++ can achieve great precision while it reveals its weakness to preserve true matches, resulting in the poor recall rate. GFC achieves the worse results due to the sensitivity to outliers. In addition, LNPG runs at a relatively slow speed with the matrix inversion computation, while it can achieve excellent matching results on the four datasets compared with the other eight competing methods.

4.4. Applications to two-view geometry estimation

The proposed LNPG is able to not only distinguish inliers from outliers, but also classify the inliers into different geometric transformation models between image pairs. Therefore, it is easy to apply the proposed LNPG to solving two-view geometry estimation problems, even in the presence of multiple structure data. We consider the two-view based motion segmentation task, which aims to segment feature matches in two views into different moving objects, and the two-view based plane segmentation task, which aims to detect the feature matches belonging to the same plane in two views. We choose the publicly available AdelaideRMF dataset [48], which contains single structure data and multiple structure data for evaluation. The performance is characterized by clustering error by using Eq. (17) as in [26,30].

4.4.1. Qualitative results

Some typical results obtained by LNPG are shown in Fig. 10. There are two, three, two, and four structures respectively, which are denoted by different colors. Note that, the intuitive matching result with only inliers on each image pair is shown in the left plot, but the motion field of all the matches is shown in the right plot. From the experimental results, we observe that the multiple structures can be effectively identified as different clusters by using LNPG, and thus they are almost segmented perfectly. This can show that the proposed LNPG is able to accurately estimate model instances from input data without requiring a user-specified number of structures.

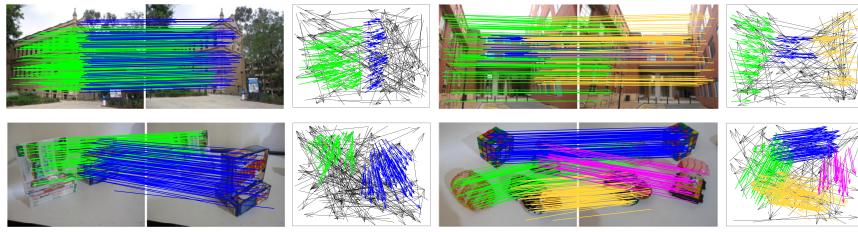


Fig. 10. Qualitative illustrations of segmentation results obtained by LNPG on the four image pairs from AdelaideRMF. From top to bottom and left to right: Oldclassicswing, Neem, Gamebiscuit, Cubebreadtoychips. For visibility, in each image pair, only inliers belonging to different structures are drawn by lines in different colors, but outliers are not shown. We also show the motion field, where the arrows corresponding to inlier clusters are masked in different colors, while the arrows corresponding to outliers are masked in the black color.

Table 2

The clustering errors (%) obtained by the six methods for two-view based motion segmentation on image pairs from the AdelaideRMF dataset. The best results are boldfaced.

Data (#)	%Outlier	RCMSA	RansaCov	MSHF	RFMSCAN	CBG	LNPG
Biscuit (1)	57.68	1.52	1.76	2.21	0.30	2.61	0.30
Book (1)	44.32	3.42	2.78	2.62	2.14	2.89	4.80
Cube (1)	69.49	4.37	3.90	3.61	1.66	3.11	3.31
Game (1)	73.48	4.21	2.92	6.35	0.86	3.61	0.43
Cubechips (2)	51.62	4.65	9.04	3.87	2.46	2.54	0.00
Cubetoy (2)	41.42	4.10	4.66	4.06	1.61	1.77	1.61
Breadcube (2)	32.19	8.93	10.02	2.73	2.89	1.74	0.41
Gamebiscuit (2)	51.54	6.77	7.44	4.24	2.13	5.18	0.30
Breadtoy (2)	37.41	6.53	17.40	24.83	1.04	6.11	1.04
Breadtoycar (3)	34.15	9.04	21.08	9.82	1.21	8.92	7.83
Biscuitbook (2)	47.51	3.75	4.91	4.90	1.17	0.59	1.76
Biscuitbookbox (3)	37.21	8.26	17.59	8.92	0.77	2.93	2.32
Breadcubechips (3)	35.22	10.09	22.30	11.91	1.74	3.30	1.74
Cubebreadtoychips (4)	28.03	6.48	24.74	16.45	3.06	13.70	2.75
Breadcartoychips (4)	35.20	14.09	30.46	21.10	2.53	6.24	2.95
Carchipscube (3)	36.59	9.58	27.82	19.64	0.61	4.24	2.42
Toycubecar (3)	36.36	11.90	34.90	25.40	7.00	12.70	8.00
Boardgame (3)	42.48	16.77	20.72	21.15	14.70	15.13	9.10
Dinobooks (3)	44.54	18.50	27.31	23.00	25.83	12.11	17.30
Total Avg.		8.05	15.36	11.41	3.88	5.76	3.60
Total Std.		4.65	10.80	8.59	6.23	4.51	4.30

(‘#’ denotes the actual number of structures in each data and ‘%Outlier’ denotes the percentage of gross outliers.)

4.4.2. Quantitative results

To conduct a quantitative evaluation for the two-view geometry estimation problems, five state-of-the-art methods for robust feature matching are adopted for comparisons, including RCMSA [25], RansaCov++ [26], MSHF [27], RFMSCAN [28] and CBG [30]. We compare LNPG with these five competing methods on the publicly available dataset (i.e., AdelaideRMF) for evaluation. We have tried our best to tune the parameters of all competing methods according to the suggestion of original papers for their excellent performance.

For two-view based motion segmentation, from the quantitative results reported in Table 2, we can see that the total average clustering error of the proposed LNPG is slightly higher than that of RFMSCAN but significantly lower than those of the other four competing methods, i.e., CBG, MSHF, RansaCov, and RCMSA. This is because that, LNPG and RFMSCAN effectively integer spatial consistency constraint into the feature matching task, but CBG, MSHF and RansaCov do not consider the constraint, improving their performance of motion segmentation. RCMSA also uses the spatial constraint to construct neighborhood graphs, but it is hard to determine appropriate neighborhood sizes for robust feature matching. RansaCov achieves the worst results due to the unreliable consensus set coverage for inlier maximization. The recently proposed CBG can achieve good performance in most data, but it fails in some data that have large number of structures, such as “Cubebreadtoychips”. Overall, the proposed LNPG can effectively separate inliers from outliers and label the inliers according to the moving objects they belong to.

For two-view based plane segmentation, from the quantitative results reported in Table 3, we can see that the proposed LNPG achieves the lowest total average clustering error among all the competing methods. The recently proposed CBG also achieves promising results, while its results are less accurate than those of LNPG. The results of RCMSA are marginally worse than that of the proposed LNPG, which is more robust to outliers. RansaCov achieves similar results with MSHF, but it needs to manually set the number of structures in the data. RFMSCAN obtains the unsatisfying results with the highest total average clustering error for two-view based plane segmentation. Overall, with the effective local neighbor preservation and propagation, LNPG achieves promising results on the most of image pairs.

Table 3

The clustering errors (%) obtained by the six methods for two-view based plane segmentation on all the other image pairs from the AdelaideRMF dataset.

Data (#)	%Outlier	RCMSA	RansaCov	MSHF	RFMSCAN	CBG	LNPG
Bonython (1)	75.13	0.51	2.68	2.98	1.52	0.00	3.54
Physics (1)	46.60	6.60	8.96	0.85	4.72	7.55	0.94
Unionhouse (1)	18.78	7.59	2.74	0.63	0.30	0.30	1.63
Elderhalla (2)	60.75	8.22	1.40	0.84	1.40	0.93	0.93
Hartley (2)	62.22	1.06	10.22	8.72	10.94	1.25	8.75
Library (2)	56.13	4.28	5.21	4.37	21.40	5.21	10.23
Sene (2)	44.49	6.53	1.48	2.00	18.40	0.72	0.80
Nese (2)	30.29	5.20	7.24	0.42	30.32	0.55	0.08
Ladysymon (2)	33.48	22.78	10.57	4.05	22.78	14.85	14.77
Oldclassicswing (2)	32.23	4.64	8.84	1.93	18.73	1.32	0.00
Neem (3)	37.83	19.00	13.58	3.36	36.93	8.60	6.64
Napiera (2)	64.73	19.54	17.30	28.87	26.16	8.74	17.22
Barrsmith (2)	69.79	9.46	23.86	34.94	12.45	10.62	9.96
Elderhallb (3)	49.80	14.67	14.90	16.43	27.84	12.08	14.51
Napierb (3)	37.13	15.14	25.14	16.45	37.84	15.06	15.06
Johnsona (4)	21.25	4.56	9.17	6.54	56.30	21.77	10.19
Unihouse (5)	18.78	15.29	8.95	15.53	66.17	15.59	7.77
Bonhall (6)	6.43	9.04	26.61	30.82	53.84	24.70	27.06
Johnsonb (7)	12.02	12.11	29.66	34.52	44.38	27.18	26.16
Total Avg.		9.80	12.03	11.28	25.92	9.32	9.28
Total Std.		6.38	8.77	12.35	19.43	8.70	8.30

(#' denotes the actual number of structures in each data and '%Outlier' denotes the percentage of gross outliers.)

5. Conclusions

In this paper, we propose a robust mismatch removal method (called LNPG) by obtaining reliably local topological structures of feature matches for removing false matches between two-view images. LNPG is able to deal with the heavily contaminated data by capturing the connection relationships of the data between different neighborhoods for effective outlier removal. Specifically, we first design a novel neighborhood graph by leveraging the spatial consistency constraint and the residual information to effectively preserve the local neighborhood structures of potential inliers. To enhance the robustness of the neighborhood graph representation, we introduce the local neighbor propagation strategy for the graph by introducing the path-based similarity measurement and the entropy-threshold-based graph partition. This strategy can expand the connection relationships of the data between different neighborhoods and thus generate a reliable neighborhood graph for mismatch removal. Based on the graph, we propose a novel consistency-filtering-based clustering algorithm to make appropriate use of the local and global relationships of feature matches for the cluster detection and cluster merging. The clustering algorithm enables the proposed LNPG to simultaneously distinguish inliers from outliers and classify the inliers into different transformation models between image pairs. Qualitative and quantitative experiments on various images with different types of image transformations have shown the superiority of the proposed method over other state-of-the-art methods.

Compared to most existing methods, LNPG is able to construct reliable neighborhoods for each feature match by taking advantage of the connection relationships of feature matches between different neighborhoods for mismatch removal. More importantly, LNPG can handle the data with multiple structures and classify the inliers into different transformation models without providing the number of structures in data. It is worth pointing out that, when input data contain a large number of feature matches, the efficiency and effectiveness of calculating the path-based similarities will be severely affected. Thus, in our future work, we shall propose a more efficient and robust method by adopting some optimization strategy to calculate similarities and handle this problem.

CRediT authorship contribution statement

Hanlin Guo: Conceptualization, Funding acquisition, Methodology, Software, Writing – original draft, Writing – review & editing. **Guobao Xiao:** Funding acquisition, Methodology, Supervision, Writing – review & editing. **Lumei Su:** Formal analysis, Writing – review & editing. **Jiaxing Zhou:** Formal analysis, Validation, Visualization. **Da-Han Wang:** Conceptualization, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 62072223, by the Scientific Research Foundation of Xiamen University of Technology under Grant YKJ22062R, by Scientific Research Foundation for Young and Middle-aged Teachers in Fujian Province JAT220342, by Natural Science Foundation of Xiamen under Grant 3502Z20227072, and by Natural Science Fund of Fujian Province under Grant 2022J05286.

References

- [1] M.S.A. Shohag, X. Zhao, Q.J. Wu, F. Pourpanah, Cross-graph reference structure based pruning and edge context information for graph matching, *Inf. Sci.* 616 (2022) 1–15.
- [2] J. Ma, J. Jiang, C. Liu, Y. Li, Feature guided Gaussian mixture model with semi-supervised em and local geometric constraint for retinal image registration, *Inf. Sci.* 417 (2017) 128–142.
- [3] E. Levinkov, A. Kardoost, B. Andres, M. Keuper, Higher-order multicut for geometric model fitting and motion segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 45 (1) (2022) 608–622.
- [4] W. Liu, C. Wang, S. Chen, X. Bian, B. Lai, X. Shen, M. Cheng, S.-H. Lai, D. Weng, J. Li, Y-net: learning domain robust feature representation for ground camera image and large-scale image-based point cloud registration, *Inf. Sci.* 581 (2021) 655–677.
- [5] C. Harris, M. Stephens, et al., A combined corner and edge detector, in: *Proc. Alv. Vis. Conf.*, 1988, pp. 10–5244.
- [6] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2) (2004) 91–110.
- [7] Y. Zheng, D. Doermann, Robust point matching for nonrigid shapes by preserving local neighborhood structures, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (4) (2006) 643–649.
- [8] C. Zhao, Z. Cao, C. Li, X. Li, J. Yang, Nm-net: mining reliable neighbors for robust feature correspondences, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2019, pp. 215–224.
- [9] J.-W. Bian, W.-Y. Lin, Y. Liu, L. Zhang, S.-K. Yeung, M.-M. Cheng, I. Reid, Gms: grid-based motion statistics for fast, ultra-robust feature correspondence, *Int. J. Comput. Vis.* 128 (6) (2020) 1580–1593.
- [10] J. Ma, A. Fan, X. Jiang, G. Xiao, Feature matching via motion-consistency driven probabilistic graphical model, *Int. J. Comput. Vis.* 130 (9) (2022) 2249–2264.
- [11] G. Wang, Y. Chen, X. Zheng, Gaussian field consensus: a robust nonparametric matching method for outlier rejection, *Pattern Recognit.* 74 (2017) 305–316.
- [12] J. Ma, J. Zhao, J. Jiang, H. Zhou, X. Guo, Locality preserving matching, *Int. J. Comput. Vis.* 127 (5) (2019) 512–531.
- [13] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, Orb: an efficient alternative to sift or surf, in: *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 2564–2571.
- [14] Y. Yao, Y. Zhang, Y. Wan, X. Liu, X. Yan, J. Li, Multi-modal remote sensing image matching considering co-occurrence filter, *IEEE Trans. Image Process.* 31 (2022) 2584–2597.
- [15] K.M. Yi, E. Trulls, Y. Ono, V. Lepetit, M. Salzmann, P. Fua, Learning to find good correspondences, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2018, pp. 2666–2674.
- [16] L. Dai, Y. Liu, J. Ma, L. Wei, T. Lai, C. Yang, R. Chen, Ms2dg-net: progressive correspondence learning via multiple sparse semantics dynamic graph, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2022, pp. 8973–8982.
- [17] L. Tang, Y. Deng, Y. Ma, J. Huang, J. Ma, Superfusion: a versatile image registration and fusion network with semantic awareness, *IEEE/CAA J. Autom. Sin.* 9 (12) (2022) 2121–2137.
- [18] X. Liu, G. Xiao, R. Chen, J. Ma, Pgfnnet: preference-guided filtering network for two-view correspondence learning, *IEEE Trans. Image Process.* 32 (2023) 1367–1378.
- [19] L. Shen, J. Zhu, C. Fan, X. Huang, T. Jin, A novel affine covariant feature mismatch removal for feature matching, *IEEE Trans. Geosci. Remote Sens.* 60 (2021) 1–13.
- [20] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* 24 (6) (1981) 381–395.
- [21] J. Ma, X. Jiang, A. Fan, J. Jiang, J. Yan, Image matching from handcrafted to deep features: a survey, *Int. J. Comput. Vis.* 129 (1) (2021) 23–79.
- [22] E. Hu, L. Sun Daniel, A fast and robust consensus maximization method for point cloud registration with high outlier ratios, *Inf. Sci.* 614 (2022) 563–579.
- [23] C.-M. Cheng, S.-H. Lai, A consensus sampling technique for fast and robust model fitting, *Pattern Recognit.* 42 (7) (2009) 1318–1329.
- [24] A. Fan, J. Ma, X. Jiang, H. Ling, Efficient deterministic search with robust loss functions for geometric model fitting, *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (11) (2021) 8212–8229.
- [25] T.T. Pham, T.-J. Chin, J. Yu, D. Suter, The random cluster model for robust geometric fitting, *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (8) (2014) 1658–1671.
- [26] L. Magri, A. Fusello, Multiple model fitting as a set coverage problem, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2016, pp. 3318–3326.
- [27] H. Wang, G. Xiao, Y. Yan, D. Suter, Searching for representative modes on hypergraphs for robust geometric model fitting, *IEEE Trans. Pattern Anal. Mach. Intell.* 41 (3) (2019) 697–711.
- [28] X. Jiang, J. Ma, J. Jiang, X. Guo, Robust feature matching using spatial clustering with heavy outliers, *IEEE Trans. Image Process.* 29 (2019) 736–746.
- [29] Q. Li, X. Lan, J. Li, Information fusion on the two-layer network for robust estimation of multiple geometric structures, *Inf. Sci.* 530 (2020) 148–166.
- [30] S. Lin, H. Luo, Y. Yan, G. Xiao, H. Wang, Co-clustering on bipartite graphs for robust model fitting, *IEEE Trans. Image Process.* 31 (2022) 6605–6620.
- [31] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2004.
- [32] L. Magri, A. Fusello, T-linkage: a continuous relaxation of j-linkage for multi-model fitting, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2014, pp. 3954–3961.
- [33] H. Wang, T.-J. Chin, D. Suter, Simultaneously fitting and segmenting multiple-structure data with outliers, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (6) (2012) 1177–1192.
- [34] M. Tepper, G. Sapiro, Nonnegative matrix underapproximation for robust multiple model fitting, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, 2017, pp. 2059–2067.
- [35] O.N. Oyelade, A.E.-S. Ezugwu, T.I. Mohamed, L. Abualigah, Ebola optimization search algorithm: a new nature-inspired metaheuristic optimization algorithm, *IEEE Access* 10 (2022) 16150–16177.
- [36] W. Zhang, D. Zhao, X. Wang, Agglomerative clustering via maximum incremental path integral, *Pattern Recognit.* 46 (11) (2013) 3056–3065.
- [37] J.O. Agushaka, A.E. Ezugwu, L. Abualigah, Dwarf mongoose optimization algorithm, *Comput. Methods Appl. Mech. Eng.* 391 (2022) 114570.
- [38] B. Zhou, X. Tang, H. Zhang, X. Wang, Measuring crowd collectiveness, *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (8) (2014) 1586–1599.
- [39] L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz, A.H. Gandomi, The arithmetic optimization algorithm, *Comput. Methods Appl. Mech. Eng.* 376 (2021) 113609.
- [40] L. Abualigah, D. Yousri, M. Abd Elaziz, A.A. Ewees, M.A. Al-Qaness, A.H. Gandomi, Aquila optimizer: a novel meta-heuristic optimization algorithm, *Comput. Ind. Eng.* 157 (2021) 107250.
- [41] L. Abualigah, M. Abd Elaziz, P. Sumari, Z.W. Geem, A.H. Gandomi, Reptile search algorithm (rsa): a nature-inspired meta-heuristic optimizer, *Expert Syst. Appl.* 191 (2022) 116158.

- [42] L. Ferraz, R. Felip, B. Martínez, X. Binefa, A density-based data reduction algorithm for robust estimators, in: Proc. Iberian Conf. Pattern Recognit. Image Anal., 2007, pp. 355–362.
- [43] Q.-H. Tran, T.-J. Chin, G. Carneiro, M.S. Brown, D. Suter, In defence of ransac for outlier rejection in deformable registration, in: Proc. Eur. Conf. Comput. Vis., 2012, pp. 274–287.
- [44] H. Liu, S. Yan, Common visual pattern discovery via spatially coherent correspondences, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2010, pp. 1609–1616.
- [45] X. Jiang, J. Jiang, A. Fan, Z. Wang, J. Ma, Multiscale locality and rank preservation for robust feature matching of remote sensing images, *IEEE Trans. Geosci. Remote Sens.* 57 (9) (2019) 6462–6472.
- [46] X. Jiang, J. Ma, A. Fan, H. Xu, G. Lin, T. Lu, X. Tian, Robust feature matching for remote sensing image registration via linear adaptive filtering, *IEEE Trans. Geosci. Remote Sens.* 59 (2) (2020) 1577–1591.
- [47] Y. Xia, J. Ma, Locality-guided global-preserving optimization for robust feature matching, *IEEE Trans. Image Process.* 31 (2022) 5093–5108.
- [48] H.S. Wong, T.-J. Chin, J. Yu, D. Suter, Dynamic and hierarchical multi-structure geometric model fitting, in: Proc. IEEE Int. Conf. Comput. Vis., 2011, pp. 1044–1051.