

Neighborhood geometry based feature matching for geostationary satellite remote sensing image



Dan Zeng^a, Ting Zhang^a, Rui Fang^a, Wei Shen^{a,*}, Qi Tian^b

^a Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, Shanghai, China

^b University of Texas at San Antonio, Texas, USA

ARTICLE INFO

Keywords:

Feature matching
Neighborhood geometry
Geostationary satellite remote sensing image
GSHHG database

ABSTRACT

In this paper, we focus on Global Self-consistent, Hierarchical, High-resolution Geography (GSHHG) database registration for remote sensing images taken from geostationary meteorological satellites. While the accuracy of feature matching is the key component. To improve it, we propose a neighborhood geometry-based feature matching scheme which includes three steps: neighborhood coding, verification and fitting. (1) Neighborhood coding represents landmarks of GSHHG as a descriptive bit-matrix, and quantifies remote sensing images to a probability-based edge map and a binary geometry-based edge map. As a result, both gradient and geometry similarity of local features in the remote sensing image and GSHHG can be measured. (2) Neighborhood verification is to encode spatial relationship among local features in neighbor, and discover outliers. (3) Neighborhood fitting fits the shorelines of GSHHG with the landmarks registered by neighborhood verification to improve recall. Experimental results on 25 pairs of newly annotated images show that the proposed method is competitive to several prior arts with respect to matching accuracy. What is more, our method is significantly more efficient than others.

1. Introduction

Edge is the region of interest in remote sensing image analysis [1,2]. Since landmarks of shoreline correspond to the edges in remote sensing images, landmark registration serves as the basis for geostationary meteorological satellite (GSMS) application. For example, it can be utilized to adjust the satellite attitude angles [3]. Fig. 1 shows the shorelines in a remote sensing image taken by a geostationary meteorological satellite and a landmark image generated by GSHHG database. The aim of landmark registration is to align the GSMS image with the landmark image. In general, landmark registration is implemented through feature points matching. However, due to particular characteristics of GSMS remote sensing images, it brings new challenges for local feature matching.

Local feature matching is a common step in many remote sensing image registration methods. Due to the invariance to affine transformation, SIFT [4] and its variants, such as SURF [5], ASIFT [6,7] and Perspective-SIFT [8], are shown to be the most successful local appearance features for remote sensing image registration [9,10]. TCSIFT and LMSIFT are applied for large-scale video copy retrieval [11] and 3D face authentication [12]. However, in GSMS images,

different landmarks tend to have similar local appearances, which inevitably lead to outliers even with SIFT-based features.

In order to address the above problem, some geometry-based methods are proposed for feature matching verification. Among these methods, Random Sample Consensus [13] (RANSAC) is the most popular one. It utilizes geometric coordinates to remove outliers [14]. Others such as Graph Transformation Matching (GTM) [15], Restricted Spatial Order Constraints (RSOC) [16] and the Triangle-area Representation of the K Nearest Neighbors (KNN-TAR) [17] are also proposed for feature matching verification. GTM and RSOC discover mismatched points according to neighborhood geometric structure, while KNN-TAR verifies feature matching based on the triangle-area representation of the K nearest neighbors.

However, due to particular characteristics of GSMS images, the previous methods have three limits: (1) SIFT based landmark registration inevitably leads to outliers. (2) The geometric similarity methods, such as RANSAC, GTM and RSOC, are extremely time consuming due to large number of pixels existing in a GSMS remote sensing image. For example, a visible image captured by FengYun-2 meteorological satellites has almost 100,000,000 pixels. (3) The constraint specified by some geometry based methods, such as KNN-TAR, is so strict that

* Corresponding author.

E-mail addresses: dzeng@shu.edu.cn (D. Zeng), tingzhang.one@gmail.com (T. Zhang), rui.f.shu@gmail.com (R. Fang), wei.shen@shu.edu.cn (W. Shen), qitian@cs.utsa.edu (Q. Tian).

<http://dx.doi.org/10.1016/j.neucom.2016.08.105>

Received 14 March 2016; Received in revised form 5 August 2016; Accepted 19 August 2016

Available online 17 November 2016

0925-2312/ © 2016 Elsevier B.V. All rights reserved.

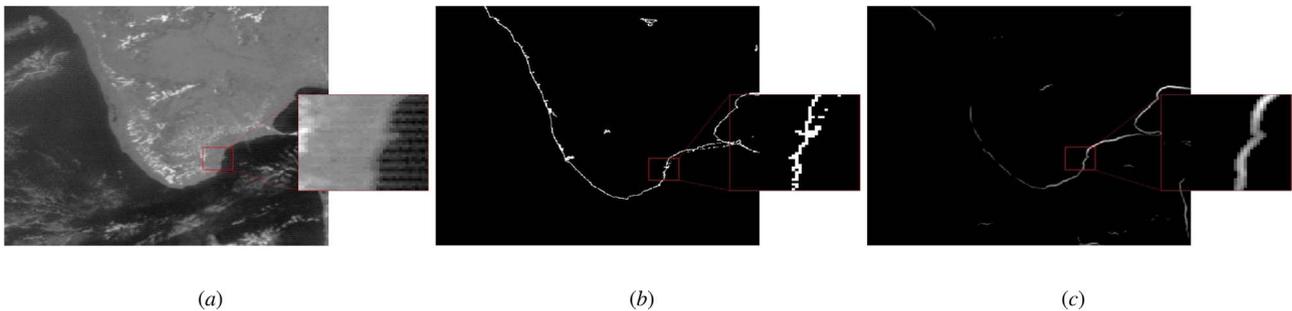


Fig. 1. Shorelines in a remote sensing image and a landmark image. (a) The geostationary meteorological satellite remote sensing image. (b) The shorelines of GSHHG database in the landmark image. (c) The edges map extracted from the GSMS image.

the recall rate is low.

To address the three problems, we propose a neighborhood geometry-based feature matching scheme including three steps: neighborhood coding, verification and fitting. Before neighborhood coding, GSHHG database and GSMS images need to be preprocessed. GSHHG database is transformed to the landmark image (as shown in Fig. 1(b)) by modeling the geo-stationary satellite image. The edges of GSMS images (as shown in Fig. 1(c)) are extracted. Neighborhood coding extracts local features from the landmark and the edge images after preprocessing. With neighborhood coding, GSHHG is described as a descriptive bit-matrix, and a GSMS image is quantized to a probability-based edge map and a binary geometry-based edge map. Since landmarks in GSHHG correspond to the edges of the GSMS images, it can simplify remote sensing image alignment and registration. We then match local features to generate initial matching pairs. As different landmarks may have similar local appearances, mismatching is inevitable. Neighborhood verification algorithm is used to remove outliers. It encodes the spatial relationships among local features in neighborhood, compares neighborhood geometric structure and detects outliers. However, neighborhood verification causes missed matching. In order to obtain more feature matching, neighborhood fitting fits the shorelines of GSHHG with the edges of remote sensing images.

The remainder of this paper is organized as follows. Section 2 gives a brief review of the related work. Section 3 introduces our approaches in details. The experimental results are presented in Section 4 and conclusions are given in Section 5.

2. Related work

Since landmarks correspond to the shorelines in remote sensing images, landmark registration is accomplished through the matching of shorelines between GSHHG and remote sensing images. Jung et al. [18] proposed a novel scheme named geocoding method. In their work, edge information in SAR data is changed to vectors, which generates candidates of image chips. GSHHG is used to generate landmarks. Then 2D correlation between image chips and landmarks is calculated. Since the shorelines are inconsistent between GSHHG and SAR data, this method does not work. Therefore, we perform landmark registration with local features.

Local feature based landmark registration consists of two parts: feature matching and feature verification.

For feature matching, SIFT is the most commonly used. Mikolajczyk et al. compared the performance of local descriptors for affine transformations, scale changes, rotation, blur, jpeg compression, and illumination changes [19]. They found that SIFT is among the best. The varieties of SIFT improve the performance further. Moreover, in order to enhance the descriptive capability of local feature, feature selection [20] and fusion [21,22] were proposed. However, since there are a lot of similar contents in GSMS remote sensing image, using intensity-based or gradient-based similarity measurement may lead to

mismatching.

For feature verification, some geometry-based algorithms are proposed to improve precision. Aguilar et al. proposed Graph Transformation Matching (GTM) [15]. It establishes a K-Nearest-Neighbor (KNN) graph to express neighbor geometric structure and determines mismatched points according to the difference of KNN graph. Liu et al. proposed Restricted Spatial Order Constraints (RSOC) algorithm [16]. RSOC designs a filtering strategy based on two-way geometric order constraints and two decision criteria restrictions. Zhang et al. proposed a triangle-area representation of the K nearest neighbors (KNN-TAR) [17]. It utilizes the descriptor KNN-TAR to find the candidate outliers and removes the real outliers by the local structure and global information. Yang et al. proposed hierarchical semantic visual coding [23–25]. Zhou et al. proposed spatial coding for large-scale partial-duplicate web image retrieval [26,27]. It takes all matching feature pairs into account, encodes their coordinates and detects mismatched pairs from them. However, the constraint specified by this method is too strict for landmark registration. Since the earth is non-standard spheroid, there is offset in global relative location. In addition to imaging principle of geo-stationary meteorological satellites, it also has position deviation phenomenon in GSMS images. Spatial relationship is consistent within local areas. Although spatial coding is much faster than other global verification methods mentioned above, it will not be fast enough because of the huge number of landmarks.

Inspired by these approaches, we propose neighborhood geometry algorithm to register GSHHG for GSMS images. Neighborhood coding extracts local features from the landmark image and the edge image. We then match local features to generate initial matching pairs. In order to improve the accuracy of feature matching, neighborhood verification is used to remove outliers. At a final touch, we add neighborhood fitting next to neighborhood verification to decrease the number of missed matching points.

3. Our approach

3.1. Neighborhood coding of GSHHG

GSHHG is a high-resolution shoreline data set merged from three databases: World Vector Shorelines (WVS), CIA World Data Bank II (WDBII) and Atlas of the Cryosphere (AC). Among them, the WVS is the basis for shorelines except for Antarctica. In this paper, we mainly utilize WVS database to generate landmark images.

The earth coordinate system is established with the coordinate origin at the center of the earth. The X -axis, Y -axis and Z -axis pass through $(90^\circ\text{E}, 0^\circ\text{N})$, $(0^\circ\text{E}, 90^\circ\text{N})$, $(0^\circ\text{E}, 0^\circ\text{N})$ respectively. Landmarks in WVS are defined as $\{L_i(\alpha_i, \gamma_i)\}_{i=1}^{N_L}$, where N_L is the number of landmarks in the earth coordinate system. Satellite position can be defined by sub-satellite point (longitude α_0 , latitude γ_0) and satellite height H . We transform the earth coordinate system to sub-satellite-point-based coordinate system by rotating α_0 degree around Y -axis, and

γ_0 degree around X -axis. As a result, landmarks are denoted as $\{L_i(\alpha_i - \alpha_0, \gamma_i - \gamma_0)\}_{i=1}^{N_L}$ in sub-satellite-point-based earth coordinate system.

With the perspective model, we can calculate the location of the standard landmark database $L_i(X_i, Y_i)$. The landmark image, whose scale is consistent to remote sensing images, is generated to represent the location of landmarks. As shown in Fig. 1(b), shorelines in the landmark image correspond to edges in the remote sensing images, which means shorelines can be used to simplify alignment and registration of remote sensing images.

We perform neighborhood coding for GSHHG landmarks. Centered at a landmark, the bit-matrix with the size of $(2K + 1) \times (2K + 1)$ pixels is extracted as the neighborhood coding map. Landmarks are quantized to descriptive bit-matrix in the landmark image. With neighborhood coding, the generated coding map can be constructed with a matrix W whose size is $(2K + 1) \times (2K + 1)$. Each element is defined as follows:

$$w_{mn} = \begin{cases} 1, & \text{if pixel } (m, n) \text{ is a landmark} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

3.2. Neighborhood coding of the GSMS image

To match the shorelines in the landmark image, the edges of the GSMS image need to be extracted. In this paper, we use Structured Forests [28] for fast edge detection, which takes advantage of the structure presented in local image patches.

The Structured Forests algorithm calculates the probability of whether a pixel is on the edge or not. The probability map can be presented with a matrix P whose size is $(2K + 1) \times (2K + 1)$. Each element denotes the probability of the pixel is an edge candidate. A threshold p_0 is set then to distinguish edge candidates from noise. We encode every pixel as a binary matrix whose size is $(2K + 1) \times (2K + 1)$, indicating the neighborhood around it, and the value of matrix element is defined as follows:

$$p'(m', n') = \begin{cases} 1, & \text{if } p'(m', n') \geq p_0 \\ 0, & \text{if } p'(m', n') < p_0 \end{cases} \quad (2)$$

3.3. Local feature matching

Neighborhood coding maps of GSHHG and GSMS image are generated in the previous section. In this section, we aim at matching features between these two sources by comparing their geometric similarity and gradient similarity.

The geometry similarity between a landmark L_i in the landmark image and a pixel $I(m', n')$ in the GSMS image can be measured as follows:

$$E_{geo}(i, m', n') = \sum_{s=1}^{2K+1} \sum_{t=1}^{2K+1} W_{s,t}^i \text{AND } P'_{s,t}, \quad (3)$$

where W represents the neighborhood coding map for GSHHG, P' denotes the binary neighborhood coding map for the GSMS image.

Similarly, the gradient similarity between a landmark L_i in the landmark image and a pixel $I(m', n')$ in the GSMS image can be calculated by:

$$E_{gra}(i, m', n') = \sum_{s=1}^{2K+1} \sum_{t=1}^{2K+1} W_{s,t}^i \times P_{s,t}, \quad (4)$$

where P denotes the probabilistic neighborhood coding map.

The number of landmarks located within the template is calculated as follows:

$$C_{geo}(i, m', n') = \sum_{s=1}^{2K+1} \sum_{t=1}^{2K+1} W_{s,t}^i. \quad (5)$$

Both geometry and gradient similarity is measured to find landmarks in remote sensing images. Local feature matching algorithm is shown in Algorithm 1.

Algorithm 1. Local feature matching.

Require: W, P, P' ; threshold r_1, r_2 (r_1 are set 0.5, r_2 are set as 0.9 based on experience)
Ensure: the best matching pixel M_i for landmark L_i in W

- 1: Given landmark L_i ;
- 2: **if** $Max\{E_{geo}(i, m', n')\} \geq r_1 \times C_{geo}(i, m', n')$ **then**
- 3: **if** $\frac{Max\{E_{geo}(i, m', n')\}}{SecondMax\{E_{geo}(i, m', n')\}} \geq r_2$ **then**
- 4: **return** $Max\{E_{geo}(i, m', n')\}$ as M_i ;
- 5: **else**
- 6: calculate E_{gra} for the two matching candidates who have bigger E_{geo} than the other;
- 7: **return** the one who gets bigger E_{gra} ;
- 8: **end if**
- 9: **else**
- 10: could not find the match pixel;
- 11: **end if**

3.4. Neighborhood verification

Suppose N matched feature pairs $P = \{(R, S)\} = \{(r_1, s_1), (r_2, s_2), \dots, (r_i, s_i), \dots, (r_N, s_N)\}$ would be obtained by local feature matching, where $R = \{r_1, r_2, \dots, r_i, \dots, r_N\}$ denotes the feature points of the landmark image and $S = \{s_1, s_2, \dots, s_i, \dots, s_N\}$ represents the corresponding feature points of the remote sensing image. Local feature matching may lead to outliers since landmarks tend to have similar local appearances in GSMS remote sensing images. In order to address this problem, we utilize neighborhood verification to remove mismatched pairs.

Geometric location relationship between each pair of features is expressed by coding their relative coordinates [26,27]. In the landmark image, the relative spatial position between each feature point along the horizontal (X -axis) and vertical (Y -axis) directions is described as two binary geometric maps, called GX_r and GY_r . Similarly, the geometric relationship between each edge point is also encoded by GX_s and GY_s in the remote sensing image. Especially, the image plane is uniformly divided into $4 \times q$ quadrants to impose stricter geometric constraints. Logical Exclusive-OR operation on GX_r and GX_s , GY_r and GY_s is performed to generate D_x and D_y which represent the difference of geometric relationship along X -axis and Y -axis, respectively.

However, there is offset in global relative location since the earth is non-standard spheroid. In addition to imaging principle of geostationary meteorological satellites, it also has position deviation phenomenon in GSMS images. Spatial relationship consistency is appropriate for local areas, not global areas. Thus we apply spatial coding to encode the geometry relationship between each matching pair and its neighboring matching pairs.

We calculate the Manhattan distance matrix D_{dis} for the point set R in a landmark image. Based on D_{dis} , M nearest neighborhood points can be sought out for each feature point, which are recorded as a neighborhood index matrix whose size is $N \times M$. With neighborhood verification, D_x and D_y size of $N \times M$ are generated to represent the difference of geometric relationship between each pair of features and its neighborhood pair of features.

The relative location relationship is correct between the i th feature point and the j th feature point only when they satisfy two conditions: $D_x(i, j) = 0$ and $D_y(i, j) = 0$. So $D_x(i, j)$ and $D_y(i, j)$ can be merged into a

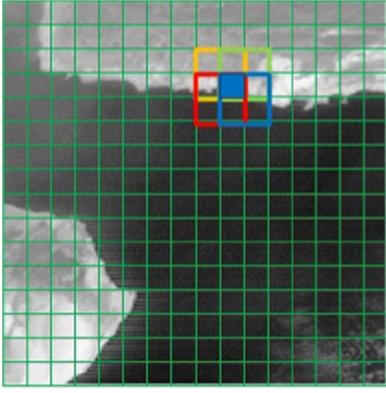


Fig. 2. Aligning shoreline with transformed patches overlaid to visualise the warp. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

single map by OR operation:

$$D(i, j) = D_x(i, j) \vee D_y(i, j). \quad (6)$$

The significance of neighborhood verification map D is whether geometric location relationship of points in the GSMS image is consistent with the landmark image. $D(i, j) = 1$ indicates the geometric relationship is inconsistent between the i th point and its j th neighbor point. On the contrary, $D(i, j) = 0$ represents spatial consistency.

The condition $D \neq 0$ indicates that some false matching pairs may exist. To confirm which feature points are mismatched, given the definition of a matrix named V as follows:

$$V(i) = \sum_{j=1}^N D(i, j). \quad (7)$$

$V(i)$ represents the number of neighborhood pairs which are inconsistent in geometry when the i th pair is selected as the reference pair. Hence the higher the value of $V(i)$, the greater possibility the i th pair is false. Based on this rule, we find the point whose V is maximum and consider this pair mismatched in the first judgment. We then remove this outlier and iterate the above process until the maximum value of V meets the threshold T .

To speed up neighborhood verification algorithm, relative location relationship D and Manhattan distance matrixes D_{dis} of all feature points are generated at the beginning of operation. Based on D_{dis} , neighborhood verification can read the neighborhood location relationship map from D .

3.5. Neighborhood fitting

Neighborhood verification removes the outliers from initial matching, which may cause missed matches. As a result, neighborhood fitting is proposed to improve recall.

The coordinate transform relationship between the feature points R in the landmark image and the feature points S in the remote sensing image is represented by the formula below:

$$\begin{pmatrix} S_x \\ S_y \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & d_x \\ 0 & 1 & d_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R_x \\ R_y \\ 1 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (8)$$

That is,

$$S(x, y) = R(x, y)H(d_x, d_y, \alpha), \quad (9)$$

where $H(d_x, d_y, \alpha)$ combines translation and rotation warp.

Neighborhood fitting splits the landmark image into small blocks. The coordinate transform relationship is considered to be consistent in a small block. In order to maintain the perspective of the blocks, each block undergoes a similarity transformation. That is, the constructed

warp H approaches a similarity transformation. To achieve this, we associate each block P_i with a cost E_i which measures the deviation of its warp function H_i from the nearest similarity transformation in the Frobenius norm, i.e.,

$$E_i(d_x, d_y, \alpha) = \min_{a_i, b_i} \iint_{(x,y) \in \Omega_i} \left\| J_i \begin{pmatrix} x & y \\ d_x & d_y \end{pmatrix} - \begin{bmatrix} a_i & -b_i \\ b_i & a_i \end{bmatrix} \right\|_F^2 dx dy, \quad (10)$$

where Ω_i is the domain of block P_i ; $J_i(x, y; d_x, d_y, \alpha)$ is the Jacobian matrix of H_i evaluated at the blocks center point.

The block of the GSMS image has different parameters combination due to the perspective projection imaging, using a basic translation and rotation warp can result in misalignment. The transformation parameters of the k th landmark L_{ijk} in the block ij , which is the i th row and the j th column block of the whole GSMS image can be determined by formula shown below:

$$H_{i,j,k} \begin{pmatrix} n & m \end{pmatrix} = \frac{\sum_{\vartheta=1}^4 \omega^\vartheta H_{i,j,k}^\vartheta(n, m)}{4}, \quad (11)$$

where $H_{i,j,k}(n, m)$ is estimated from the weighted problem; (n, m) represents the landmark located in row n , column m of the GSMS image. As shown in Fig. 2, the blue overlapping region shares the four neighbor blocks transformation parameters information.

The scalar weights $\{\omega^\vartheta\}_{\vartheta=1}^4$ give great importance to data of the block containing initial registration areas that are close to block which has the initial matching results, the value ϑ shows that the blocks contain initial registration result in the neighborhood. The weights are calculated as:

$$\omega^\vartheta = \exp(-\|x_* - x_i\|^2 / \sigma^2), \quad (12)$$

where σ is a scale parameter; x_* is the center point of the overlapping region that performs translation and rotation warp; x_i is rest center point in the neighborhood blocks containing initial registration points.

4. Experimental result

All the experiments are performed on a workstation with dual Intel Xeon CPU (2.1 GHz and 12 cores for each) and 128 GB RAM, using Matlab (64 bit).

4.1. Data set

The critical content of our experiment is to match the landmarks in GSHHG with edges in remote sensing images. Specifically, we focus on GSMS images acquired by Fengyun-2D meteorological satellites as our remote sensing images. Sub-satellite point of Fengyun-2D is around (86 °E, 0 °N), and only landmarks located within $\pm 60^\circ$ of longitude and $\pm 60^\circ$ of latitude around sub-satellite point are chosen as testing data, concerning radial distortion.

The size of GSMS images is 10000×10000 pixels. Taking efficiency into consideration, we divide GSMS images into patches of 400×400 pixels to match their respective feature points. Some of the shorelines can not be detected in GSMS images due to the occlusion of clouds, resulting in difficulties of matching these shorelines during local feature matching process. So we select 25 patches with relatively more edges in GSMS images in the process of removing outliers.

To measure matching performance, we manually labels the ground truth: for each landmark in the landmark image, we accurately mark its corresponding point in the GSMS images.

4.2. Neighborhood coding

For local feature matching, the value of K is essential since it determines the size of feature template. As a result, we validate the

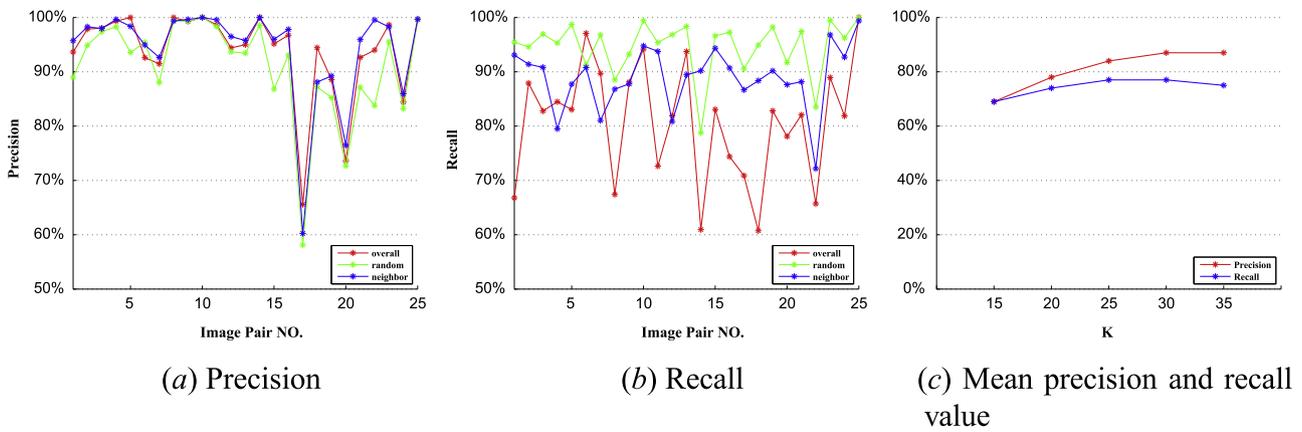


Fig. 3. Performance of local feature matching with different K ($(2K + 1) \times (2K + 1)$ is the size of feature template). The smaller K value is, the lower precision is. The larger K value is, the lower recall is. To trade off precision and recall, K is set as 30.

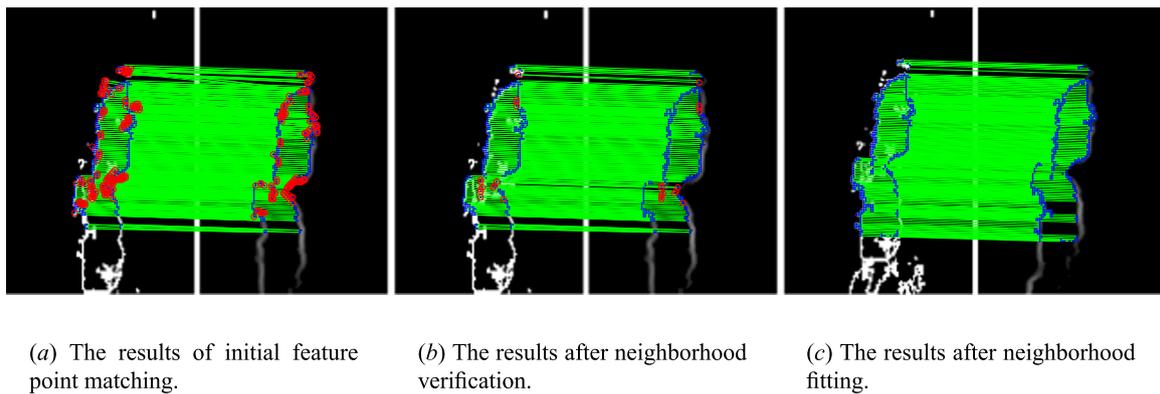


Fig. 4. The comparison of results after neighborhood coding, verification and fitting. The points surrounded by red circles is mismatching points. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

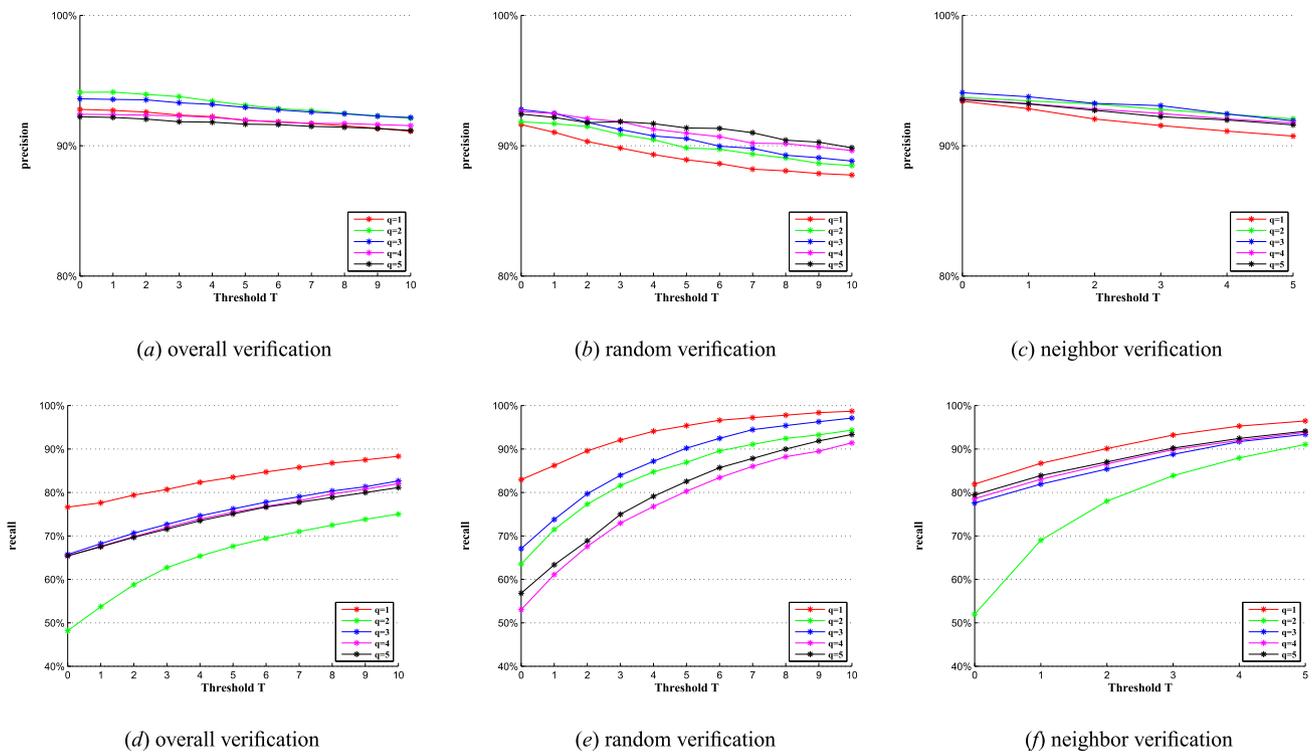


Fig. 5. Mean recall and precision values of overall, random and neighborhood verification with different q (the number of coordinate system) and T (the number of tolerated error points). The larger value of q or smaller value of T signifies the higher precision and lower recall. To trade off precision and recall, q is set as 3. For overall, random and neighborhood verification, T is set as 8, 7, 3 respectively.

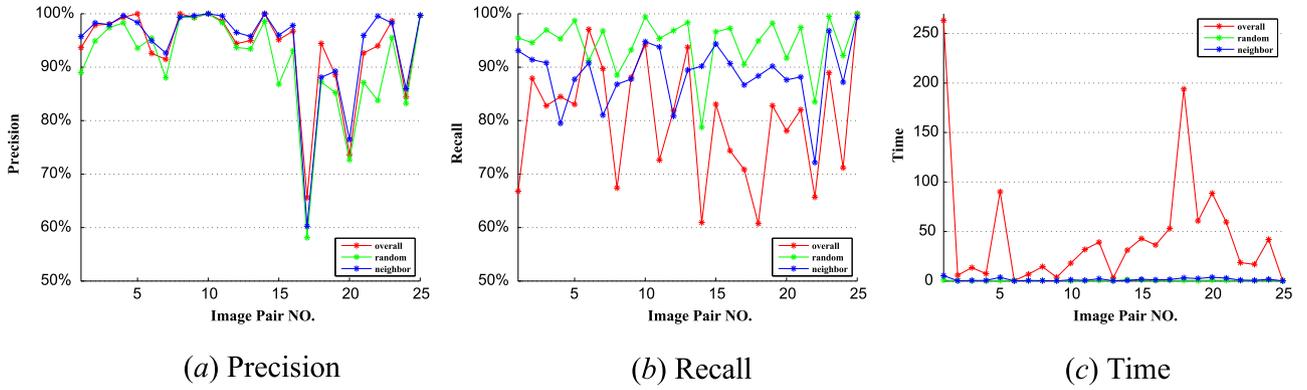


Fig. 6. Performance comparison of overall, random and neighbor verification. Neighborhood verification precedes overall and random verification.

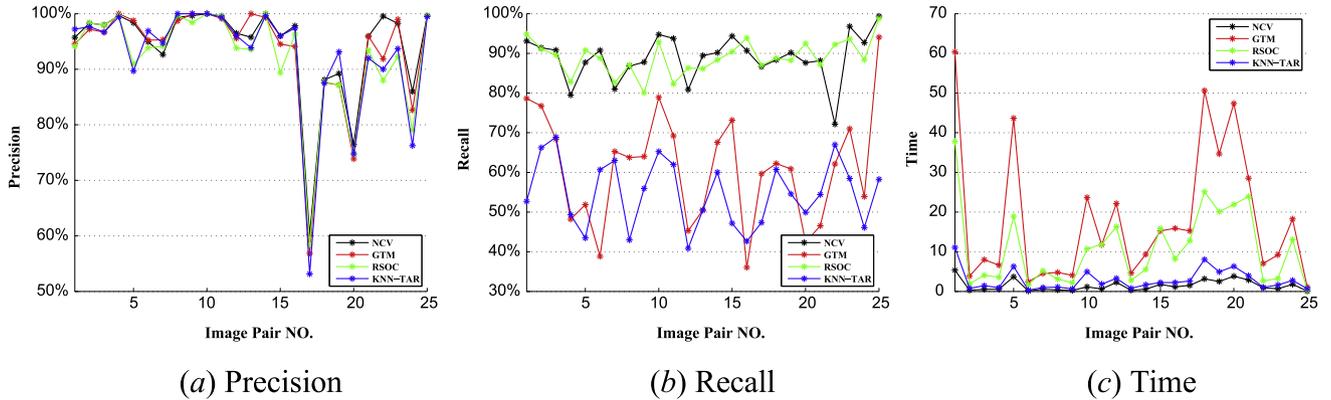


Fig. 7. Performance of three algorithms on 25 image sketch pairs. NCV is competitive to GTM, RSOC and KNN-TAR in precision and recall. NCV significantly reduces time cost.

Table 1

Mean recall, precision and time values in NCV, GTM, RSOC and KNN-TAR (Baseline is the initial matching results.).

	NCV	GTM	RSOC	KNN – TAR
Precision (%)	95.7	94.9	93.6	94.4
Recall (%)	88.9	61.2	90.0	54.9
Time (s)	1.48	18.12	10.92	2.91

value of K while fixing the value of r_1 and r_2 to 0.5 and 0.9 respectively. As shown in Fig. 3, K varies from 15 to 35 at a step size of 5. Fig. 3(a) shows the precision value for selected 25 patches. And Fig. 3(b) expresses the recall value. Fig. 3(c) gives mean precision and recall value to observe the effect of the value of K on performance.

From the above results, the value of K is set to 30. In other words, the size of template is set to $61 \times 61 (K = 30)$ for GSMS images of 10000×10000 pixels. Fig. 4(a) is the results of initial matching. There are mismatched feature points (points in red circles as shown in Fig. 4(a)) after local feature matching due to the error of feature extraction caused by Structured Forests and interruption of clouds.

4.3. Neighborhood verification

For geometric verification, we adopt three strategies: overall, random and neighborhood verification. Overall verification means that all of feature points are verified together ($M=N$). In random verification, feature pairs are grouped randomly and every group is verified respectively. Neighborhood verification considers every feature point with its neighboring feature points. There are two parameters, q and T , needing to be set.

In order to get proper values of q and T , three different coding schemes are applied to 25 image patches in Fig. 5. Considering the tradeoff between recall and precision, the threshold T is set to 8, 7 and

3 in overall, random and neighborhood verification respectively. By comparing the performance of results when $q=1$, $q=2$, $q=3$, $q=4$ and $q=5$, we set the value of q as 3 in these three geometric verification algorithms.

To select the best geometric verification methods, we compare these three methods in case they get the optimal parameters. Fig. 6 shows the performance comparison. Fig. 6(a) reveals that the precision of neighborhood verification is the highest and the precision of random verification is the lowest, while as shown in Fig. 6(b), the recall value of random verification is the highest. Neighbor verification takes second place. Fig. 6(c) reveals that the time cost for random and neighborhood verification is close, while overall verification is much more time consuming. Taking all these factors into consideration, we choose neighborhood verification to make the tradeoff between the precision and recall values. After neighborhood verification, the result of feature matching is shown in Fig. 4(b).

4.4. Comparison among feature matching algorithms

Fig. 7 presents the statistical results of four algorithms: Neighborhood Coding and Verification (NCV), Graph Transformation Matching (GTM), Restricted Spatial Order Constraints (RSOC) and the Triangle-area Representation of the K Nearest Neighbors (KNN-TAR) on 25 image patches. Table 1 gives the mean of above results. As shown in Fig. 7(a), the precision of these four algorithms is close in general. However, Table 1 indicates the average precision of NCV is slightly higher compared with GTM, RSOC and KNN-TAR. Fig. 7(b) indicates that the recall values of NCV and RSOC are close. Since GTM and KNN-TAR are too strict with spatial location relationship among feature points, their recall values are relatively low. As shown in Fig. 7(c), NCV significantly outperforms the three other algorithms with respect to time efficiency.

In addition, NCV algorithm has the advantage of low time complex-

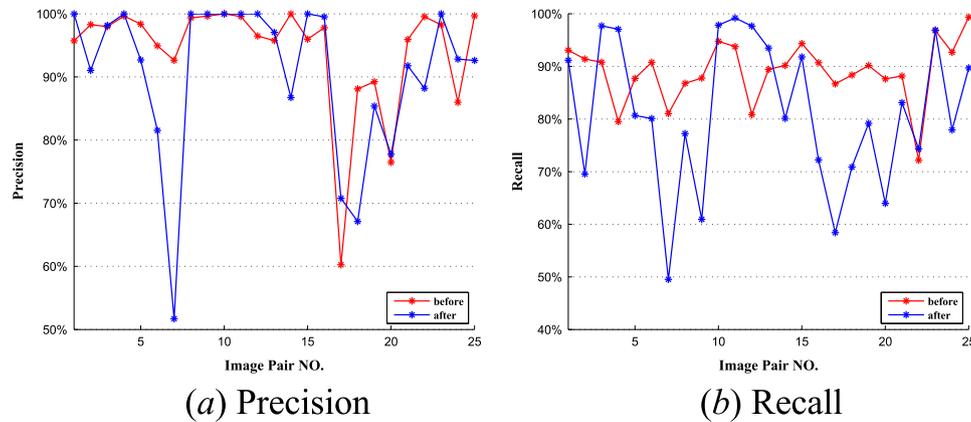


Fig. 8. Performance comparison before and after neighborhood fitting. The values of precision are close, but the recall rate after neighborhood fitting significant increases.

Table 2

Mean recall and precision values before and after neighborhood fitting. (Baseline is the ground truth.).

	Before	After
Precision (%)	95.7	91.3
Recall (%)	70.2	81.9

ity. Assuming that there would be N feature pairs in initial matching results, we analyze time complexity of NCV, GTM, RSOC and KNN-TAR in the worst case of $N_{outliers} = N$. The computation time complexity of GTM, RSOC and KNN-TAR is $O(N^3)$. Since NCV codes K neighboring points around the reference point, its computation time complexity is $K \times O(N^2) = O(N^2)$. Only when $K=N$, the computation time complexity of NCV is same as GTM, RSOC and KNN-TAR, and neighbor verification is equal to overall verification in this situation.

As mentioned above, NCV algorithm is competitive to GTM, RSOC and KNN-TAR in the precision and recall, and significantly reduces time cost.

4.5. Neighborhood fitting

In order to improve recall, we fit the shorelines consisting of landmarks after neighborhood verification. As shown in Fig. 8, the precision and recall are compared before and after neighborhood fitting. The values of precision are close, but the recall after neighborhood fitting increases. Especially, the precision of 7th image pair has a sharp drop. Due to cloud covering, the visible shorelines are short and straight, which leads to aperture effect. Landmarks are fitted with cloud edges. This issue can be solved with cloud detection [29]. Table 2 gives mean recall and precision values before and after neighborhood fitting.

Fig. 4(c) shows the result of neighborhood fitting. Compared with 4(b), the number of feature points increases and there are no mismatched points. Thus it can be seen that neighborhood fitting can not only improve the recall rate, but can also further refine the feature points.

5. Conclusion

In this paper, we implement feature matching between the shorelines of GSHHG and the edges of GSMS images by neighborhood geometry algorithms.

We utilize edge information as a stable matching feature in GSMS images. Neighborhood coding describes GSHHG and GSMS images with a landmark map and an edge probability map respectively. It aims to extract features as many as possible and achieve the high recall value.

In order to improve the precision, neighborhood verification algorithm utilizes spatial location relationship of features distribution to remove outliers. It includes three strategies: overall, random and neighbor verification. Experimental results show the performance of neighborhood verification is the best among the three. And compared with GTM, RSOC and KNN-TAR, neighborhood verification also has the advantage in the precision, recall and time cost.

In order to improve the recall, neighborhood fitting algorithm fits the shorelines of GSHHG and the edges of GSMS remote sensing images. Based on neighborhood topological structure, neighborhood fitting is able to match these features missed in the neighborhood verification. Compared with other methods, the recall rate significantly increases.

However, neighborhood geometry algorithm cannot be invariant to rotation, therefore we mainly apply it to feature matching in GSMS images. In the following work, we will further improve this algorithm for developing its application in other remote sensing images.

Acknowledgments

This work is supported by National Natural Science Foundation of China (61301221 and 61572307).

References

- [1] L. Zhang, J. Chen, B. Qiu, Region of interest extraction in remote sensing images by saliency analysis with the normal directional lifting wavelet transform, *Neurocomputing* 179 (2016) 186–201.
- [2] W. Guang-hui, W. Shu-bi, W. Hua-bin, L. Can-hai, T. Xin-ming, T. Jiao-jiao, T. Juan, An algorithm of parameters adaptive scale-invariant feature for high precision matching of multi-source remote sensing image, in: *Urban Remote Sensing Event, 2009 Joint, IEEE, 2009*, pp. 1–7.
- [3] F. Tang, X. Zou, H. Yang, F. Weng, Estimation and correction of geolocation errors in fengyun-3c microwave radiation imager data, *IEEE Trans. Geosci. Remote Sens.* 54 (1) (2016) 407–420.
- [4] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2) (2004) 91–110.
- [5] A. Brook, E. Ben-Dor, Automatic registration of airborne and spaceborne images by topology map matching with surf processor algorithm, *Remote Sens.* 3 (1) (2011) 65–82.
- [6] G. Yu, J.-M. Morel, A fully affine invariant image comparison method, in: *IEEE International Conference on Acoustics, Speech and Signal Processing, 2009, ICASSP 2009, IEEE, 2009*, pp. 1597–1600.
- [7] S. Cao, J. Jiang, G. Zhang, Y. Yuan, An edge-based scale-and affine-invariant algorithm for remote sensing image registration, *Int. J. Remote Sens.* 34 (7) (2013) 2301–2326.
- [8] G.-R. Cai, P.-M. Jodoin, S.-Z. Li, Y.-D. Wu, S.-Z. Su, Z.-K. Huang, Perspective-sift: an efficient tool for low-altitude remote sensing image registration, *Signal Process.* 93 (11) (2013) 3088–3110.
- [9] H. Goncalves, L. Corte-Real, J.A. Goncalves, Automatic image registration through image segmentation and sift, *IEEE Trans. Geosci. Remote Sens.* 49 (7) (2011) 2589–2600.
- [10] B. Fan, C. Huo, C. Pan, Q. Kong, Registration of optical and sar satellite images by exploring the spatial relationship of the improved sift, *IEEE Geosci. Remote Sens. Lett.* 10 (4) (2013) 657–661.
- [11] Y. Zhu, X. Huang, Q. Huang, Q. Tian, Large-scale video copy retrieval with

- temporal-concentration sift, *Neurocomputing* 187 (4) (2016) 83–91.
- [12] Y. Ming, X. Hong, A unified 3d face authentication framework based on robust local mesh sift feature, *Neurocomputing* 184 (4) (2016) 117–130.
- [13] 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.
- [14] P. Qiu, Y. Liang, The improved algorithm of remote sensing image registration based on sift and contourlet transform, in: 2013 4th IEEE International Conference on Software Engineering and Service Science (ICSESS), IEEE, 2013, pp. 906–909.
- [15] W. Aguilar, Y. Frauel, F. Escolano, M.E. Martinez-Perez, A. Espinosa-Romero, M.A. Lozano, A robust graph transformation matching for non-rigid registration, *Image Vis. Comput.* 27 (7) (2009) 897–910.
- [16] Z. Liu, J. An, Y. Jing, A simple and robust feature point matching algorithm based on restricted spatial order constraints for aerial image registration, *IEEE Trans. Geosci. Remote Sens.* 50 (2) (2012) 514–527.
- [17] K. Zhang, X. Li, J. Zhang, A robust point-matching algorithm for remote sensing image registration, *IEEE Geosci. Remote Sens. Lett.* 11 (2) (2014) 469–473.
- [18] J.S. Jung, J.H. Song, Y.K. Kwag, High precision automatic sar geocoding method using gshhs for coastal ship monitoring, in: 9th European Conference on Synthetic Aperture Radar, 2012, EUSAR, VDE, 2012, pp. 360–363.
- [19] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (10) (2005) 1615–1630.
- [20] L. Zheng, S. Wang, Z. Liu, Q. Tian, Fast image retrieval: query pruning and early termination, *IEEE Trans. Multimedia* 17 (5) (2015) 648–659.
- [21] L. Zheng, S. Wang, Q. Tian, Coupled binary embedding for large-scale image retrieval, *IEEE Trans. Image Process.* 23 (8) (2014) 3368–3380.
- [22] L. Zheng, S. Wang, Z. Liu, Q. Tian, Packing and padding: Coupled multi-index for accurate image retrieval, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1939–1946.
- [23] Y. Yang, H. Zhang, M. Zhang, F. Shen, X. Li, Visual coding in a semantic hierarchy, in: Proceedings of the 23rd Annual ACM Conference on Multimedia Conference, ACM, 2015, pp. 59–68.
- [24] Y. Yang, Z.-J. Zha, Y. Gao, X. Zhu, T.-S. Chua, Exploiting web images for semantic video indexing via robust sample-specific loss, *IEEE Trans. Multimedia* 16 (6) (2014) 1677–1689.
- [25] Y. Yang, Z. Ma, Y. Yang, F. Nie, H.T. Shen, Multitask spectral clustering by exploring intertask correlation, *IEEE Trans. Cybern.* 45 (5) (2015) 1083–1094.
- [26] W. Zhou, Y. Lu, H. Li, Y. Song, Q. Tian, Spatial coding for large scale partial-duplicate web image search, in: Proceedings of the 18th ACM International Conference on Multimedia, ACM, 2010, pp. 511–520.
- [27] W. Zhou, H. Li, Y. Lu, Q. Tian, Large scale image search with geometric coding, in: Proceedings of the 19th ACM International Conference on Multimedia, ACM, 2011, pp. 1349–1352.
- [28] P. Dollár, C.L. Zitnick, Structured forests for fast edge detection, in: 2013 IEEE International Conference on Computer Vision (ICCV), IEEE, 2013, pp. 1841–1848.
- [29] B.B. Barnes, C. Hu, A hybrid cloud detection algorithm to improve modis sea surface temperature data quality and coverage over the eastern Gulf of Mexico, *IEEE Trans. Geosci. Remote Sens.* 51 (6) (2013) 3273–3285.



Dan Zeng received the B.E. and Ph.D degrees from University of Science and Technology of China (USTC) in 2003 and 2008, respectively. She is currently an associate professor in the Department of Computer Engineering at Shanghai University, and a visiting scholar at the University of Texas at San Antonio (UTSA). Her research interests include computer vision and image/video analysis.



Ting Zhang received the B.E. in communication and information engineering from Shanghai University, Shanghai, China, in 2014. She is now a master's candidate in Shanghai University. Her research interests focus on local feature matching and image registration, especially for remote sensing images.



Rui Fang received his B.E. in communication and information engineering from Shanghai University, China, in 2014. He is pursuing his Master's Degree in Shanghai University now. His main topics of research are computer vision and image registration, especially for geostationary satellite images.



Wei Shen received his B.S. and Ph.D. degrees both in Electronics and Information Engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2007 and in 2012, respectively. From April 2011 to November 2011, he worked in Microsoft Research Asia as an intern. Now he is currently an assistant Professor with the School of Communication and Information Engineering, Shanghai University, Shanghai, China. His research interests include shape analysis, edge and symmetry detection and human pose estimation.



Qi Tian received the B.E. in electronic engineering from Tsinghua University, China, in 1992, the M.S. in electrical and computer engineering from Drexel University, Philadelphia, PA, in 1996, and the Ph.D in electrical and computer engineering from the University of Illinois, Urbana-Champaign, Champaign, IL, in 2002. He is currently a professor in the Department of Computer Science at the University of Texas at San Antonio (UTSA), San Antonio, TX. He took a one-year faculty leave at Microsoft Research Asia (MSRA) during 2008–2009. His research interests include multimedia information retrieval and computer vision. He has published over 280 refereed journal and conference papers. His research projects were funded by NSF, ARO, DHS, SALSI, CIAS, and UTSA and he also received faculty research awards from Google, NEC Laboratories of America, FXPAL, Akiira Media Systems, and HP Labs. He received the Best Paper Awards in PCM 2013, MMM 2013 and ICIMCS 2012, the Top 10% Paper Award in MMSP 2011, the Best Student Paper in ICASSP 2006, and the Best Paper Candidate in PCM 2007.