Multisource high-resolution optical remote sensing image registration based on point–line spatial geometric information

Heng Yan,a Shuwen Yang,a,b,c,* Yikun Li,a,b Qing Xue,a and Mengsheng Zhanga

aLanzhou Jiaotong University, Faculty of Geomatics, Lanzhou, China
bNational-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring, Lanzhou, China
cGansu Province Engineering Laboratory for National Geographic State Monitoring, Lanzhou, China

Abstract. High-resolution multisource optical remote sensing images often have considerable non-linear radiation and scale differences that become more prominent as the image resolution increases. Since this problem may cause low registration accuracy, we propose an automatic registration algorithm based on point–line spatial geometric information (PLSGI). First, we propose an improved scale invariant feature transform algorithm using the quadtree uniformization, the Bhattacharyya distance, and the slope constraint to obtain many correct point pairs. Then we generate a new global line descriptor using PLSGI. In particular, the descriptor can resist the non-linear radiation and scale differences of multisource remote sensing images. Finally, we use a piecewise linear model to warp the sensed images based on a set of tie points that consist of corresponding features and corresponding intersections. The effectiveness of the proposed algorithm was validated using multiple groups of high-resolution remote sensing images. The experimental results indicated that our algorithm is highly generalized and robust for high-resolution multisource remotely sensed images. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JRS.15.036520]

Keywords: multisource image registration; high-resolution satellite image; spatial geometric feature; point–line.

Paper 210160 received Mar. 16, 2021; accepted for publication Sep. 9, 2021; published online Sep. 22, 2021.

1 Introduction

Since single-source images often provide insufficient information, it is becoming a trend in remote sensing applications to integrate the complementary information of multisource high-resolution remote sensing images. Image registration is the process of transforming two or more images into the same coordinate system, which facilitates the comprehensive utilization of information contained in the matched images from different sensors, viewpoints, resolutions, and time periods.1 Its accuracy considerably influences the validity of subsequent image processing and applications, such as image fusion,2,3 change detection,4,5 three-dimensional (3D) reconstruction,6,7 and remote sensing quantitative information analysis.8

High-spatial-resolution remote sensing images have distinct layers, clear texture, and rich spatial information; thus they have broad applications.9 Multisource optical remote sensing images are often very different in non-linear radiation and scale. Moreover, they have rich and detailed texture information in different ground types such as shadows, roads, and rivers. However, these characteristics significantly interfere with the registration process and thus pose considerable challenges to classical registration methods, which can be essentially partitioned into two categories: area-based10–13 and feature-based methods.14–18 Feature-based methods can overcome the shortcomings of area-based methods, such as high calculation demand, sensitivity

*Address all correspondence to Shuwen Yang, yangshuwen@mail.lzjtu.cn

to intensity changes, and partial occlusion. Therefore, feature-based methods have become mainstream image registration methods. In addition, considering the inconsistent deformation caused by terrain relief and degradation factors, many scholars have adopted new robust registration strategies. For example, Lee\textsuperscript{19} proposed a coarse-to-fine approach for register multispectral remote sensing images and remove outliers using restricted spatial order constraints. Feng et al.\textsuperscript{20,21} introduced an automatic remote sensing mountain image registration method based on optical flow estimation and proposed another algorithm that combines and locates feature-based and area-based methods to achieve robust registration.

The point feature is the most widely used (point, line, edge, contour, and region) in image registration. In this category, the SIFT is the most classic method; it highly resists scale, noise, partial occlusion, and illumination variations between to-be-matched images.\textsuperscript{22} Scholars have made tremendous efforts to improve this method. For example, the OS-SIFT resists the geometric and intensity differences between high-resolution optical and SAR images and improves the registration accuracy.\textsuperscript{23} Moreover, the UR-SIFT is a robust point descriptor that acquires evenly distributed point features in multisource remotely sensed images with significant illumination, rotation, and scene differences.\textsuperscript{24} However, these point-based methods may fail to deal with specific multisource high-resolution remote sensing images, which have unstable gradient information caused by large differences in non-linear radiation and scale.

Compared with the point feature, the line feature within an image provides rich semantic information and enforces stronger constraints on the spatial structure. Therefore, more scholars have developed methods based on line features in recent years and have obtained some vital research results.\textsuperscript{25,26} The mean-standard deviation line descriptor (MSLD) is a line descriptor that has robust rotation, illumination, and viewpoint variants based on local gradient information.\textsuperscript{27,28} The line-point invariants (LPI) calculates the similarity of line features through the affine invariance, which effectively improves the discrimination of line features in the matching process.\textsuperscript{29} The line-junction-line (LJL) was proposed to increase the number of correct matches (NCM) in remote sensing images.\textsuperscript{30} The above line-based methods are effective in certain cases. However, they are not robust enough for large-scale and non-linear radiation differences, which are common characteristics of multisource high-resolution remote sensing image.

Compared with natural images with stable distribution and uniform structure of image units, terrain changes, viewpoint changes, and cloud or shadow occlusion may cause local geometric distortion in remote sensing images, increasing the difficulty of registration. In particular, the imaging conditions of multisource HR optical remote sensing images are heterogeneous and complex, which may lead to the inability to extract precisely located corresponding features under the influence of various unstable factors such as atmospheric medium, temperature, humidity, and radiation intensity. Therefore, to solve the above problems, we propose an automatic registration algorithm based on point–line spatial geometric information (PLSGI). This research has three significant contributions: first, an improved SIFT method combines the quad-tree uniformization, the Bhattacharyya distance, and the slope constraint to acquire more correct point matches. Second, a spatial geometric relationship-based global line descriptor is designed based on PLSGI. Third, corresponding features are used together with corresponding intersections to accurately estimate the transformation model. The experimental results show that our algorithm has higher generalization and registration accuracy when dealing with multisource HR remote sensing images compared with other related algorithms. Moreover, it is robust to scale and illumination variations.

The organization of the remainder of this paper is as follows. The details of our improved SIFT method and the line feature matching method by spatial geometric information are given in Secs. 2 and 3, respectively. Section 4 describes the proposed registration algorithm for multisource HR optical remote sensing images in detail. After that, we provide the experimental comparisons and discussion in Sec. 5. Finally, Sec. 6 concludes this paper.

## 2 Improved SIFT Method

The number of corresponding points on the remote sensing images will affect the robustness of image registration, and the standard SIFT methods do not acquire enough stable corresponding
points on multisource high-resolution remote sensing images. So we propose an improved SIFT method for extracting corresponding points. Our improvements are in the three ways: feature point extraction based on quadtree uniformization, coarse matching, and outlier elimination.

First, the Harris detector is used to extract the feature points, and their uniform distribution is achieved using the quadtree algorithm. If the distribution of feature points in space is more hierarchical and uniform, the feature matching will represent a spatial geometric relationship more accurately. Figure 1 compares original feature point extraction using the Harris detector and the feature point uniformization based on the quadtree algorithm. To reduce redundancy, the feature point with the highest response value is selected from each node as the unique feature point for this node, and all other low-response value feature points within the node are removed. Therefore, feature point uniform extraction can provide a balance between the uniformity of distribution of corresponding points and computational complexity.

Then the Bhattacharyya distance is used as a similarity measure instead of the Euclidean distance to finish coarse match, as the Euclidean distance between feature vectors is usually determined by the larger bin value while often ignoring some smaller bin values, which may cause some mismatches. The Bhattacharyya distance is calculated according to the following equation:

$$D_B(X, Y) = \sum_{i=1}^{n} \sqrt{x_i y_i},$$

where $X$ and $Y$ are the two normalized feature vectors.

Figure 2 illustrates the result of coarse matching for panchromatic image data using the Bhattacharyya distance as a similarity measure and a distance ratio threshold set to 0.9. The image data are from the GaoFen-1 (GF-1) and GaoFen-2 (GF-2) satellites with corresponding resolutions of 1 and 2 m. Although severe non-linear radiation and large-scale differences exist...
between the two panchromatic images tested in Fig. 2(a), many initially matched points pairs were still obtained since the Bhattacharyya distance is sensitive to the feature vector difference. Figure 2(b) indicates that all slopes of the initially matched point pairs for the test images are within \([-1.0, 2.5]\), which contain some outliers. Their slopes are significantly different from the slope of correct matches. Consequently, it is effective and convenient to use the difference between the slopes of feature point pairs to eliminate outliers, provided that the initially matched point pairs contain most correct matches.

In the process of fine matching, purifying the matching point pairs with the slope constraint can retain an appropriate NCM while eliminating outliers. The remotely sensed images perform global correction based on the RPC files prior to application, which can largely eliminate distortions in the remote sensing images. In this case, the basic assumption for applying the slope constraint is that all correct matching point pairs will have similar slope values under the visual system. This means that the correct matching point pairs have a strong aggregation. The slope of the matched point pairs is calculated by

\[
s_i = \frac{y_{i2} - y_{i1}}{(x_{i2} + w) - x_{i1}},
\]  

where \((x_{i1}, y_{i1})\) and \((x_{i2}, y_{i2})\) denote the coordinates of the \(i\)'th matching point pair in the reference and sensed images, respectively, and \(w\) denotes the horizontal displacement distance. The steps of the slope constraint algorithm are as follows.

1. The slope of all initial point pairs is calculated, and the slope histogram is constructed.
2. The feature point pairs are grouped according to the slope histogram.
3. The location error is iteratively calculated for each group of feature points, and the point pairs with the largest residuals are eliminated until the precision is satisfied.
4. Each group of precisely located feature point pairs constitutes a corresponding point.

3 Line Feature Matching by Spatial Geometric Information

3.1 Multiscale Line Feature Extraction

The line feature is a stable image feature that not only resists the non-linear radiation differences between images but also has rich semantic information. The line detection algorithms proposed in the past are prone to deviating the edge feature’s direction when influenced by noise with large gradients or intersection of the line features. This will cause a longer line to be cut into several short lines. Thus it is difficult to apply these line detection algorithms to high-resolution remote sensing images with complex and detailed information. The CannyLines is a line detector free of parameters and with accuracy location, which is well suited for extracting line features in meter and submeter resolution remote sensing images. This algorithm is composed of four main parts: (1) edge map extraction using the parameter-free CannyPF algorithm; (2) edge linking and splitting; (3) line feature extension and merging; and (4) line verification based on the Helmholtz principle.

The performance of the CannyLines, EDLines, line segment detector (LSD), fast line detector (FLD), and progressive probabilistic hough transform (PPHT) was tested on a GF-2 remote sensing image [Fig. 3(a)] and is shown in Fig. 3(b)-(f). The results reveal that the CannyLines most effectively extracted accurately located clear line features that truly reflected the original line structure in the remote sensing image [Fig. 3(b)]. The major causes for the high performance of line detection using the CannyLines can be summarized from two perspectives: one is that it detects line features directly from an edge map, which can avoid losing some useful information when detecting line features from the edge features; the other is that all line features are validated using the Helmholtz principle, which considers both gradient direction and magnitude information. The detection results of the EDLines [Fig. 3(c)], LSD [Fig. 3(d)], and FLD [Fig. 3(e)] were almost as good, but they extracted broken and redundant line features. Moreover, some line features tended to deviate from their actual positions. The detection results of the PPHT [Fig. 3(f)] were unsatisfactory, and they were all scattered short-line features.
Therefore, we detected line features on HR remote sensing images by the CannyLines to obtain more reliable line features.

To detect line features on multisource high-resolution remote sensing images that have large-scale differences, subsampling is utilized to build multiscale spatial pyramids used by the standard SIFT method. The multiscale space in this study is a Gaussian pyramid with no inner layer between the two successive octaves. The steps of line feature extraction in multiscale space are as follows.

First, calculate the Gaussian scale space \( L(x, y, s) \) by Eq. (3), which is computed by convolution of the image \( I(i, j) \) and the Gaussian function \( G(i, j, \sigma_s) \):

\[
L(i, j, s) = G(i, j, \sigma_s) \ast I(i, j),
\]

(3)

where \( \sigma_s \) denotes the scale space factor of the \( s \)'th octave image, which reflects the blurring effect.

Then the CannyLines algorithm is used to detect initial line features in each individual scale of the image. To avoid too many extracted unstable line features and involved calculations, we eliminate the unstable line features to obtain stable line features. For a line feature \( l_i \) at the \( s \)'th octave image, the orientation is determined by the gradient orientation of each pixel point on it, which is defined as

\[
\theta(x, y, s) = \tan^{-1}\left[\frac{L(x+1, y+1, s) - L(x, y-1, s)}{L(x+1, y+1, s) - L(x-1, y, s)}\right].
\]

(4)

We quantize the gradient orientation by Eq. (5) into \( T \) bins to facilitate the calculation:

\[
\hat{\theta} = \text{mod}\left(\frac{\theta(x, y, s)}{2\pi/T} + \frac{1}{2}, T\right).
\]

(5)

Finally, the orientation histogram \( R_i \) of the line feature \( l_i \) is established using Eq. (6), and the orientation corresponding to its peak is determined to be the main orientation of line feature \( l_i \). If more than 50% of the scales have the same orientation, this line feature \( l_i \) is stable and preserved:

\[
R_i = \sum \delta(\hat{\theta}^j, t) \quad (t = 0, 1, T-1) \quad \delta(a, b) = \begin{cases} 
1, & a = b \\
0, & a \neq b
\end{cases}
\]

where \( \sigma_s \) denotes the scale space factor of the \( s \)'th octave image, which reflects the blurring effect.
3.2 Line Feature Description and Matching

Many of the current line feature descriptors are based on neighborhood gradient information, which cannot provide stable features when there are large radiation differences between images, especially non-linear radiation differences. PLSGI refers to the relative geometric relationship between line features and the Delaunay triangulation constructed by corresponding points; as the corresponding points have one-to-one correspondence on the reference and sensed images, the Delaunay triangulation built by these points is unique, in which each edge of the Delaunay triangulation on the reference and sensed images also has a unique spatial correspondence. Accordingly, a novel line feature descriptor is generated by PLSGI, which can overcome the large non-linear radiation differences between HR remote sensing images. The robustness of our descriptor is directly determined by the uniform distribution of corresponding points in the remote sensing imagery. The complete generation process of proposed descriptor is shown in Fig. 4.

First, the Delaunay triangulation is built using corresponding points obtained by the improved SIFT method and the logarithmic polar coordinate grid is established with the main direction of the line feature as the reference direction, which is centered on the midpoint of the line feature.

The global region is divided into $N$ rings, and each ring has the same width. Moreover, the $u$’th ring $(u = 1,...,N)$ is divided into $M_u$ subregions, and the $v$’th subregion $(v = 1,...,M_u)$ of this ring is denoted by $R(u,v)$. The angular quantization number of the outer ring should be larger than the inner ring to ensure that the size of each subregion is relatively uniform.

Second, gain the spatial geometric information $DOL$ between a line feature $l_i$ and each edge $e_j$ of $DT$ is obtained. $DOL$ consists of three components:

(a) the distance ratio $D_{i,j}$ between $l_i$ and $e_j$;
(b) the angle $O_{i,j}$ between $l_i$ and $e_j$; and
(c) the length ratio $L_{i,j}$ between $l_i$ and $e_j$.

![Fig. 4](https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing on 19 Nov 2021 Terms of Use: https://www.spiedigitallibrary.org/terms-of-use)
The spatial geometric information \( DOL \) is calculated via the following equation:

\[
D_{i,j} = \frac{d_1(i,j)}{d_2(i,j)}, \quad O_{i,j} = \theta_{i,j}, \quad L_{i,j} = \frac{|l_i|}{|e_j|},
\]  

(7)

where \( d_1(i,j) \) and \( d_2(i,j) \) denote the distances between \( l_i \) and two endpoints of \( e_j \), respectively; \(|l_i|\) and \(|e_j|\) denote the length of \( l_i \) and \( e_j \), respectively; and \( \theta_{i,j} \) indicates the angle between \( l_i \) and \( e_j \).

The distance ratio \( D_{i,j} \) and the angle \( O_{i,j} \) can accurately represent the spatial relative position between the line feature and the Delaunay triangulation, and the length ratio \( L_{i,j} \) is used to limit the lengths of the line matches to be approximately close. Therefore, each unique line feature can be described in sufficient detail using the above spatial geometric information and is not affected by the non-linear radiative differences between high-resolution remote sensing images.

Then the cumulative values of \( DOL \) are calculated by Eq. (8) to generate histogram \( h(u, v) \) in each subregion. Each edge falling in the subregion is recorded in \( DT \), and its location attribute is represented by its midpoints:

\[
h(u, v) = \{h^d, h^o, h^l\},
\]  

(8)

where \( h^d \), \( h^o \), and \( h^l \) denote the cumulative values of \( D_{i,j} \), \( O_{i,j} \), and \( L_{i,j} \) in the \( v \)th subregion of the \( u \)th ring, respectively. They are calculated by

\[
h^d = \sum_{j=1}^{N} C_i; j, \quad h^o = \sum_{j=1}^{N} O_i; j, \quad h^l = \sum_{j=1}^{N} L_i; j.
\]  

(9)

Finally, the Gaussian function \( w_u \) is used to weigh \( h(u, v) \) via Eq. (10) to improve the robustness of our descriptor against image geometric distortions. The subregions close to the descriptor center are assigned a larger weight value, subregions farther from the descriptor center are assigned a smaller weight value, and all subregions within the same circle have the same weight:

\[
h^l(u, v) = h(u, v) \cdot w(u) \quad w(u) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2}{2\sigma^2}}.
\]  

(10)

The three-bin vectors of each subregion are combined using Eq. (11) into \( M \)-bin \( (M = \sum_{u=1}^{N} M_u) \) feature vector \( v_i \):

\[
v_i = \{h^l(1,1), h^l(1,2), \ldots, h^l(N, M_N)\}.
\]  

(11)

To obtain more correct line matches, the cost function \( C_{i1, j2} \) is defined as in Eq. (12). A smaller \( C_{i1, j2} \) value indicates that the line features \( l_{i1} \) and \( l_{j2} \) in the reference and sensed images is more reliable as the corresponding line feature pair:

\[
C_{i1, j2} = \sqrt{\frac{1}{M} \sum_{k=1}^{M} |v_{i1}(k) - v_{j2}(k)|}.
\]  

(12)

4 Multisource High-Resolution Optical Remote Image Registration

The advantages of fully combining point and line features are considered to provide accurate and robust registration model parameters. In this section, an HR remote sensing image registration algorithm based on point–line spatial geometric relationship is introduced in detail. The flowchart of our algorithm is shown in Fig. 5; it is divided into following three steps.

1. **The acquisition of corresponding point features.** To improve the accuracy and number of corresponding points, many correct corresponding feature points are obtained using the improved SIFT method and are used to build the Delaunay triangulation.
(2) The acquisition of corresponding line features and their intersections. Multiscale line features are extracted using the CannyLines algorithm, and the main direction of the line features are calculated at each scale, while retaining the stable line features accordingly. Then the line feature descriptor based on PLSGI is generated to achieve fine matching. Subsequently, inspired by existing research, four constraints are used to enhance the accuracy of the intersection of line features:

1) Region constraint. The line feature to be searched must be contained within the search region of the reference line feature. The search area is constructed as shown in Fig. 6.

2) Angle constraint. The angle between the two line features must be >40 deg.

3) Distance constraint. The shortest distance between the intersection and all end-points of two line features must be less than the length of the corresponding line feature.

4) Cross constraint. If the intersection in the reference image and the corresponding intersection in the sensed image both satisfy the above three constraints, the two intersections are considered to be a fine matching pair; otherwise, they are ignored.

Fig. 6 The search region.
(3) **Outliers elimination and image warping.** The fast sample consensus (FSC) algorithm\(^3^8\) improves the accuracy of the results through the selection of correct matches and the elimination of imprecise matches and can obtain more correct matches in fewer iterations than the RANSAC algorithm. Therefore, the FSC algorithm is used to remove outliers. In addition, the corresponding intersection and corresponding feature points are combined into a set of tie points and a piecewise linear model is employed as the transformation model. This model divides the image into triangular regions by triangulating the filtered tie points, and the global mapping function is obtained by piecing together the affine transformation of each local region.

## 5 Experiments and Discussion

In this section, we validate the effectiveness of the proposed PLSGI algorithm through three experiments, which mainly include the performance of the improved SIFT method, the performance of line feature matching, and the registration for multisource HR remote sensing image. In particular, the effectiveness of proposed algorithm is evaluated through a combination of subjective and objective methods. The objective approach is the evaluation criteria, and the subjective approach observes the registration effect by means of the checkerboard mosaic images and the enlarged subregion maps.

### 5.1 Experiment 1: Performance of the Improved SIFT Method

#### 5.1.1 Experiment details

In this experiment, four sets of HR remote sensing image pairs from different sensors and different scenes were used; These image pairs were preregistered using RPC information. Therefore, no obvious rotation or translation differences are apparent on the test image pairs, but they still have large intensity and scale changes owing to variations in sensors and spectra. The complete information of tested image pairs is given in Table 1.

Our improved SIFT method was evaluated and compared with four other typical methods: oriented least square matching (OLSM),\(^3^9\) affine-SIFT (ASIFT),\(^4^0\) position scale orientation (PSO)-SIFT,\(^4^1\) and oriented fast and rotated BRIEF (ORB).\(^4^2\) The OLSM is an accurate affine invariant image matching method that can effectively overcome local affine features’ low positional accuracy. The ASIFT algorithm is completely affine invariant to the six parameters of the affine transformation. The PSO-SIFT algorithm performs feature matching by combining a new gradient calculation approach with feature point position. The ORB algorithm is a high-speed feature matching method with rotation invariance and noise resistance. The ORB is available in the third-party library OpenCV, and the others were implemented based on the authors’ codes.

<table>
<thead>
<tr>
<th>Data no.</th>
<th>Location</th>
<th>Size</th>
<th>GSD (m)</th>
<th>Date</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tianjin, China</td>
<td>2400 × 2400</td>
<td>1.2</td>
<td>2018/09</td>
<td>Google Earth Band 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2200 × 2200</td>
<td>1.0</td>
<td>2019/09</td>
<td>GF-2 Pan</td>
</tr>
<tr>
<td>2</td>
<td>Yining, China</td>
<td>751 × 751</td>
<td>2.1</td>
<td>2019/03</td>
<td>ZY-3 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>700 × 700</td>
<td>2.0</td>
<td>2019/07</td>
<td>Google Earth Band 1</td>
</tr>
<tr>
<td>3</td>
<td>Shenzhen, China</td>
<td>1000 × 1000</td>
<td>1.0</td>
<td>2018/08</td>
<td>GF-2 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>930 × 930</td>
<td>1.2</td>
<td>2019/04</td>
<td>Google Earth Band 1</td>
</tr>
<tr>
<td>4</td>
<td>Guangzhou, China</td>
<td>3400 × 3400</td>
<td>0.6</td>
<td>2016/09</td>
<td>Google Earth Band 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2100 × 2100</td>
<td>1.0</td>
<td>2018/12</td>
<td>GF-2 Pan</td>
</tr>
</tbody>
</table>

---

Yan et al.: Multisource high-resolution optical remote sensing image registration…


Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing on 19 Nov 2021
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
The performance of these methods is evaluated using four metrics: precision, NCM, root-mean-square error (RMSE), and distribution quality (DQ).

Equation (13) is used to calculate the precision, which reflects the percentage of correct matches in the total number of matches:

$$\text{precision} = \frac{\text{NCM}}{\text{NCM} + \text{NFM}},$$

where NCM and NFM denote the number of correct and false matches in the final matching result, respectively.

The RMSE of total matches is calculated as shown in Eq. (14); it is computed by the transformation model estimated by the correct matches:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (p_i^r - p_i)^2},$$

where \(N\) denotes the number of evenly distributed corresponding point pairs, which were manually selected from the reference and sensed images by experts in related fields.

The DQ of corresponding points is measured by the area and shape of the formed triangles. It is calculated as shown in the following equation:

$$\text{DA} = \sqrt{\frac{\sum_{i=1}^{m} \left(\frac{A_i}{\bar{A}} - 1\right)^2}{m - 1}}, \quad \bar{A} = \frac{\sum_{i=1}^{m} A_i}{m}$$

$$\text{DS} = \sqrt{\frac{\sum_{i=1}^{m} (S_i - 1)^2}{m - 1}}, \quad S_i = \frac{3 \cdot \max(J_i)}{\pi}$$

$$\text{DQ} = \text{DA} \cdot \text{DS} = \sqrt{\frac{\sum_{i=1}^{m} \left(\frac{A_i}{\bar{A}} - 1\right)^2 \cdot \sum_{i=1}^{m} (S_i - 1)^2}{m - 1}},$$

where \(m\) is the number of triangles and \(A_i\) and \(\max(J_i)\) represent the area and the largest internal angle of the \(i\)th triangle, respectively. The smaller value of DQ indicates a more uniform distribution of tie points.

### 5.1.2 Experimental results

The matching performance of the different methods is quantitatively compared, and the results are shown in Table 2. It can be clearly observed that the ORB method takes the least time, but this method is not robust to multisource HR remote sensing images with large intensity and scale differences. As a result, this method is not robust to the matching task and achieves the worst matching performance. The OLSM, ASIFT, and PSO-SIFT methods obtain higher precision and RMSE. Nonetheless, they take more time and only obtain a limited NCM, which directly results in a low DQ of corresponding points. The problem may be caused by many mistakenly rejected correct matches in the feature matching stage. The improved SIFT method achieved the best performance in the above four criteria and is only inferior to the ORB method in terms of time cost.

Figure 7 shows the resulted images and their locally enlarged regions matched by the proposed improved SIFT method. All four sets of locally enlarged regions illustrate that the corresponding point pairs are accurately positioned with minuscule deviation. The phenomenon is mainly attributed to the quadtree-based uniformization algorithm, which retains the feature points with the highest response value within each node and removes other feature points. Thus a uniform distribution of feature points is ensured with the increased NCM. Since data 1 contains many regular buildings with well-defined boundaries, the proposed algorithm obtains the maximum NCM and has the highest DQ. It is verified through data 4, which has large-scale differences, that our method is still robust to scale change and thus gains superior DQ.
Table 2  The matching performance of different methods.

<table>
<thead>
<tr>
<th>Data no.</th>
<th>Criterion</th>
<th>OLSM</th>
<th>ASIFT</th>
<th>PSO-SIFT</th>
<th>ORB</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DQ</td>
<td>2.78</td>
<td>2.44</td>
<td>2.63</td>
<td>3.16</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>39</td>
<td>50</td>
<td>48</td>
<td>14</td>
<td>278</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>74.36</td>
<td>87.71</td>
<td>72.73</td>
<td>42.68</td>
<td>91.5</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>0.93</td>
<td>1.14</td>
<td>0.67</td>
<td>2.72</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>40.51</td>
<td>23.20</td>
<td>65.16</td>
<td>2.72</td>
<td>19.96</td>
</tr>
<tr>
<td>2</td>
<td>DQ</td>
<td>1.93</td>
<td>1.42</td>
<td>1.55</td>
<td>2.57</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>56</td>
<td>72</td>
<td>65</td>
<td>37</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>69.64</td>
<td>75.79</td>
<td>81.20</td>
<td>58.48</td>
<td>85.88</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>0.81</td>
<td>1.38</td>
<td>0.64</td>
<td>1.91</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>50.01</td>
<td>25.07</td>
<td>210.29</td>
<td>1.34</td>
<td>13.23</td>
</tr>
<tr>
<td>3</td>
<td>DQ</td>
<td>2.57</td>
<td>1.78</td>
<td>2.11</td>
<td>3.03</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>35</td>
<td>65</td>
<td>54</td>
<td>22</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>74.82</td>
<td>71.47</td>
<td>83.86</td>
<td>62.23</td>
<td>89.68</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>1.03</td>
<td>1.57</td>
<td>0.70</td>
<td>2.45</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>26.89</td>
<td>17.45</td>
<td>43.62</td>
<td>1.38</td>
<td>12.23</td>
</tr>
<tr>
<td>4</td>
<td>DQ</td>
<td>1.40</td>
<td>2.23</td>
<td>2.61</td>
<td>2.32</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>152</td>
<td>133</td>
<td>71</td>
<td>42</td>
<td>248</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>76.97</td>
<td>83.12</td>
<td>68.22</td>
<td>55.81</td>
<td>86.39</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>1.34</td>
<td>1.36</td>
<td>1.59</td>
<td>1.98</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>20.29</td>
<td>12.23</td>
<td>12.84</td>
<td>3.75</td>
<td>6.13</td>
</tr>
</tbody>
</table>

Fig. 7 (a)–(d) The matching results of our improved SIFT method.
However, due to local geometric distortions, the right region of data 2 missed the corresponding point pair, resulting in a lower DQ. Since data 3 contains a large mountain region, performing feature matching is challenging, but our method still obtains a reasonable NCM. Therefore, future work will focus on designing a point descriptor robust to local geometric distortion of the remotely sensed imagery that can accommodate complex scenes such as mountainous scenes.

5.2 Experiment 2: Performance of Line Feature Matching

5.2.1 Experiment details

To validate the line matching performance of the PLSGI algorithm, three other state-of-the-art line matching algorithms including the were selected for the comparison: LPI, LJL, and MSLD; these excel in matching line features on images captured by a single sensor. All of three algorithms were implemented using the codes provided by their authors.

As scale and illumination variations tend to interfere with feature matching, we performed scale and illumination transformation on a panchromatic image with a spatial resolution of 2.1 m from the ZY-3 satellite as shown in Fig. 8 and validated the matching performance of the above three algorithms. In our experiment, the scale ratio ranged from 0.4 to 1.6 and in steps of 0.2. In addition, non-linear illumination transformation was performed by adding the same increment for each channel. The corresponding range of the illumination increments was $[-90, 90]$ and in steps of 30. Both scale and illumination transformation are implemented based on the third-party library OpenCV. The original image was used as the reference image, and an image after scale or illumination transformation was used as the sensed image in our experiments. In addition, three criteria, recall, precision, and NCM, were used to evaluate the matching performance of different algorithms.

5.2.2 Experimental results

The experimental results of the four algorithms applied to the simulated image are shown in Fig. 9. The LJL algorithm achieved the largest NCM and was very robust against illumination changes. However, its recall and precision performance were unsatisfactory under the different scale conditions. The poor performance of the LPI algorithm for line feature matching results in an inability to obtain enough correct matches, especially in the case of large-scale or large illumination variations. The MSLD algorithm was no longer robust under the scale and illumination variations, only obtaining a few correct matches. By comparison, the proposed algorithm remained highly robust to scale and illumination changes. The encouraging aspect is that our algorithm consistently maintains 100% precision of line feature matching for all scale and illumination variations, probably because multiscale line features obtained by the CannyLines algorithm can adapt to scale variations, whereas the proposed line feature descriptor based on spatial geometric information is not affected by scale and illumination changes. The average recall and NCM of the proposed method applied to simulated images obtained by a series of scale and illumination transformations were 88.9% and 857, respectively.

![Fig. 8 The example of simulated images: (a) a panchromatic image from the ZY-3 satellite; (b) scale transformation; and (c) illumination transformation.](image-url)
Figure 10 shows a part of the line feature matching results by the PLSGI algorithm for simulated images with scale and illumination variations. We marked the corresponding line feature using the same color. As can be seen from Figs. 10(a) and 10(b), our algorithm is very robust to changes of illumination and can obtain many corresponding line features in both strong and weak illumination conditions. Meanwhile, these corresponding line features can be uniformly distributed on the images and have an exact position. Figures 10(c) and 10(d) indicate that the proposed algorithm is invariant with respect to scale change. Therefore, with large scale differences, the proposed algorithm is still able to obtain a large number of corresponding line features, which are fitted to corresponding intersections to register images.

5.3 Experiment 3: Registration for Multisource HR Remote Image

5.3.1 Experiment details

In this experiment, six pairs of multisource remote sensing images were selected to analyze the availability of the proposed PLSGI algorithm as shown in Fig. 11. The upper row is the sensed
image, and the lower row is the corresponding reference image. The details of tested image pairs are shown in Table 3. These images cover different scenes including urban, factory, and suburban areas, and with spatial resolutions at the meter or even submeter level. The scale difference between the reference and sensed images in some of the data is more than twice as large, which is a challenge for image registration. Similarly, obvious rotation and translation differences were removed for all HR remote sensing images using RPC information.

The five algorithms, LPI, LJL, MSLD, ASIFT, and OLSM, were compared with the PLSGI algorithm. For the registration algorithm based on line feature, we use the intersection fitting method used in this paper to obtain the corresponding point pairs, and the FSC algorithm was also used to remove outliers to obtain the final tie points. The sensed images were warped using the piecewise linear transformation model according to the tie points obtained by the different algorithms.

### 5.3.2 Experimental results

As shown in Table 4, the MSLD algorithm only succeeds in registering data 3–4. The overall registration performance of this algorithm is the worst of the five algorithms. As the MSLD algorithm is based on the gradient information in the neighborhood region for line feature.

---

**Table 3** The complete information of the images used in experiment 3.

<table>
<thead>
<tr>
<th>Data no.</th>
<th>Location</th>
<th>Size</th>
<th>GSD (m)</th>
<th>Date</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yining, China</td>
<td>1000 × 1000</td>
<td>1.0</td>
<td>2019/09</td>
<td>GF-2 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2150 × 2150</td>
<td>0.6</td>
<td>2018/09</td>
<td>Google Earth</td>
</tr>
<tr>
<td>2</td>
<td>Guangzhou, China</td>
<td>1400 × 1400</td>
<td>1.0</td>
<td>2015/03</td>
<td>GF-2 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1300 × 1300</td>
<td>1.2</td>
<td>2015/10</td>
<td>Google Earth</td>
</tr>
<tr>
<td>3</td>
<td>Yining, China</td>
<td>1200 × 1200</td>
<td>2.1</td>
<td>2019/03</td>
<td>ZY-3 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1600 × 1600</td>
<td>2.0</td>
<td>2019/07</td>
<td>GF-1 Pan</td>
</tr>
<tr>
<td>4</td>
<td>Bozhou, China</td>
<td>1400 × 1400</td>
<td>2.1</td>
<td>2019/09</td>
<td>ZY-3 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>950 × 950</td>
<td>4.8</td>
<td>2017/08</td>
<td>Google Earth</td>
</tr>
<tr>
<td>5</td>
<td>Yining, China</td>
<td>1000 × 1000</td>
<td>2.0</td>
<td>2019/06</td>
<td>GF-1 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1500 × 1500</td>
<td>1.2</td>
<td>2020/03</td>
<td>Google Earth</td>
</tr>
<tr>
<td>6</td>
<td>Lanzhou, China</td>
<td>500 × 500</td>
<td>2.0</td>
<td>2019/09</td>
<td>GF-1 Pan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1600 × 1600</td>
<td>1.0</td>
<td>2017/05</td>
<td>GF-2 Pan</td>
</tr>
</tbody>
</table>
description and large radiation differences between multisource HR remote sensing images cause unstable gradient information, this algorithm cannot obtain a reasonable amount of line matching pairs, which ultimately lead to a failure of image registration. The LJL algorithm is effective for data 1–4 and obtains many correct matches and a high DQ, but it is deficient in precision and

Table 4  The registration performance of different algorithms.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Data no.</td>
<td>Criterion</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>LPI</td>
<td>DQ</td>
<td>1.39</td>
<td>1.93</td>
<td>2.32</td>
<td>2.85</td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>379</td>
<td>120</td>
<td>103</td>
<td>73</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>75.05</td>
<td>87.59</td>
<td>70.06</td>
<td>68.22</td>
<td>37.9</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>1.89</td>
<td>2.21</td>
<td>2.15</td>
<td>2.63</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>423.18</td>
<td>98.23</td>
<td>130.01</td>
<td>93.45</td>
<td>311.71</td>
</tr>
<tr>
<td>LJL</td>
<td>DQ</td>
<td>1.49</td>
<td>1.36</td>
<td>1.56</td>
<td>0.92</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>771</td>
<td>650</td>
<td>624</td>
<td>475</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>56.27</td>
<td>71.74</td>
<td>64.46</td>
<td>66.52</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>1.63</td>
<td>2.07</td>
<td>1.31</td>
<td>1.75</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>748.04</td>
<td>392.21</td>
<td>420.85</td>
<td>233.90</td>
<td>—</td>
</tr>
<tr>
<td>MSLD</td>
<td>DQ</td>
<td>—</td>
<td>3.25</td>
<td>2.99</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>—</td>
<td>32</td>
<td>51</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>—</td>
<td>70.96</td>
<td>76.12</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>—</td>
<td>4.46</td>
<td>1.84</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>—</td>
<td>44.05</td>
<td>54.89</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ASIFT</td>
<td>DQ</td>
<td>1.83</td>
<td>1.44</td>
<td>1.77</td>
<td>2.42</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>126</td>
<td>160</td>
<td>193</td>
<td>209</td>
<td>279</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>72.17</td>
<td>67.41</td>
<td>78.81</td>
<td>81.80</td>
<td>52.77</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>2.31</td>
<td>1.52</td>
<td>1.69</td>
<td>2.52</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>21.64</td>
<td>13.04</td>
<td>15.53</td>
<td>17.15</td>
<td>21.35</td>
</tr>
<tr>
<td>OLSM</td>
<td>DQ</td>
<td>2.13</td>
<td>2.37</td>
<td>1.78</td>
<td>1.75</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>157</td>
<td>209</td>
<td>180</td>
<td>241</td>
<td>302</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>69.23</td>
<td>78.85</td>
<td>75.56</td>
<td>81.24</td>
<td>80.02</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>2.03</td>
<td>1.66</td>
<td>1.24</td>
<td>1.73</td>
<td>2.48</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>40.95</td>
<td>31.69</td>
<td>37.81</td>
<td>31.88</td>
<td>34.67</td>
</tr>
<tr>
<td>PLSGI</td>
<td>DQ</td>
<td>0.69</td>
<td>0.71</td>
<td>0.62</td>
<td>1.16</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>NCM</td>
<td>9.46</td>
<td>605</td>
<td>713</td>
<td>371</td>
<td>399</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>93.97</td>
<td>100</td>
<td>95.85</td>
<td>100</td>
<td>93.18</td>
</tr>
<tr>
<td></td>
<td>RMSE (pixel)</td>
<td>0.75</td>
<td>0.80</td>
<td>0.92</td>
<td>0.87</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>32.97</td>
<td>21.65</td>
<td>22.963</td>
<td>19.77</td>
<td>20.05</td>
</tr>
</tbody>
</table>
RMSE while incurring a high time cost. For data 5–6 with scale differences close to a factor of 2, the LJL algorithm fails even more severely. Although the LPI algorithm completes all of the registration experiments, its RMSE is large; for example, it has an unacceptable value of 3.44 for data 6. In short, the registration performance of the algorithms based on line features only is poor. Similarly, the ASIFT and OLSM fail to achieve better registration performance for point feature-based registration algorithms due to the large-scale non-linear radiation between multisource remote sensing images, while their RMSE and DQ remain poor. Compared with the other five algorithms, the proposed PLSGI is a registration algorithm that uses both point and line features, enabling the two features to complement each other. Our algorithm always ensures that sufficient corresponding pairs are available, and even for data 6 with the twice scale difference, 304 pairs of corresponding pairs are still obtained with good DQ. The average DQ, NCM, precision, and RMSE obtained by the PLSGI algorithm for all six data sets are 0.97, 556, 95.22, and 0.96, respectively. Moreover, our algorithm is ranked second in terms of time cost.

Figure 12 shows the tie point pairs of the proposed algorithm applied to HR remote sensing images (the red points represent the point matches obtained by the improved SIFT method, and the green points represent the fitted intersection matches obtained by the corresponding line features). The scale difference in data 1–3 is relatively small, but the non-linear radiation differences are large, and the PLSGI algorithm is still able to obtain uniformly distributed tie points, which provides an advantage for high-accuracy registration. For data 4–5, the corresponding points obtained by the improved SIFT and the corresponding intersections fitted by the corresponding line feature complement each other to still obtain a uniform distribution of tie points under the large-scale difference.

Figure 13 illustrates the Delaunay triangulations constructed by tie points, which is used to warp the sensed image using piecewise linear functions. The tie point consists of corresponding feature points and corresponding intersections. It is evident that the tie points obtained by the PLSGI algorithm can cover almost the entire sensed image, thus ensuring the accuracy of the registration. Figure 14 shows the checkerboard mosaicked images of the registration results by different algorithms for six pairs of multisource HR remote sensing images. Although there are
Fig. 13 The Delaunay triangulations constructed by tie points: (a)–(f) multisource HR remote sensing image pairs no. 1–6, respectively.

Fig. 14 The mosaic results of different algorithms for multisource remote sensing image pairs. (a)–(f) Multisource HR remote sensing image pairs no. 1–6, respectively.
large scale and non-linear radiation differences, the locally enlarged regions show that the neighboring subregions still have good continuity in each checkerboard mosaicked image, which can directly reflect the effectiveness and accuracy of the proposed algorithm. It also demonstrates the robustness of the proposed algorithm to scale and non-linear radiation differences.

6 Conclusions
In this paper, we proposed an automatic image registration algorithm based on PLSGI, which is strongly robust to the differences of non-linear radiation and scale between multisource HR remote sensing images. The proposed algorithm has three components: (1) the acquisition of corresponding point features; (2) the acquisition of corresponding line features and their intersections; and (3) the outliers elimination and image warping. Our algorithm not only increases a large number of corresponding features but also provides a highly accurate registration accuracy. The experimental results on six pairs of multisource and multiscale optical images indicated that the proposed algorithm is more efficient than state-of-the-art and classic algorithms and it can achieve high-precision registration results. In the future, we will focus on improving the speed of image registration and coupling with points and curve segments to enhance the registration performance for more complex scenes.

Acknowledgments
This work was supported by grants from the National Key Research and Development Program of China (No. 2017YFB0504201) and the National Natural Science Foundation of China (No. 41761082). No conflicts of interest exist in the submission of this manuscript, and the manuscript is approved by all authors for publication.

References

Heng Yan is currently working toward his MS degree in surveying and mapping from Lanzhou Jiaotong University, Lanzhou, China. His research focuses on image processing and pattern recognition.

Shuwen Yang received his BS degree from China University of Geosciences, Wuhan, China, in 1999 and his MS degree and PhD from the School of Earth Sciences, China University of Geosciences, Wuhan, China, in 2004 and 2011, respectively. Since 2004, he has been at Lanzhou Jiaotong University, where he is currently a professor in the faculty of geomatics. His research focuses on image processing and pattern recognition.

Yikun Li received his BS degree from Xi’an Jiaotong University, Xi’an, China, in 2000 and his PhD from the School of Computer Engineering, Nanyang Technological University, Singapore, in 2008. Since 2009, he has been at Lanzhou Jiaotong University, where he is currently an associate professor in the faculty of geomatics. His research focuses on remote sensing image retrieval and change detection.

Qing Xue is currently working toward his MS degree in surveying and mapping from Lanzhou Jiaotong University, Lanzhou, China. His research focus on remote sensing image processing and analysis.

Mengsheng Zhang is currently working toward his MS degree in surveying and mapping from Lanzhou Jiaotong University, Lanzhou, China. His research focus on remote sensing image processing and analysis.