ISPRS Journal of Photogrammetry and Remote Sensing 151 (2019) 15–26

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Robust registration for remote sensing images by combining and localizing feature- and area-based methods

Ruitao Feng[a], Qingyun Du[a,c,d], Xinghua Li[b,*], Huanfeng Shen[a,c,d,*]

[a] School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China
[b] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
[c] Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China
[d] Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Keywords:
Block-weighted model
Huber estimation
Outlier-insensitive model
Registration
Structure tensor
Remote sensing image

ABSTRACT

Highly accurate registration is one of the essential requirements for numerous applications of remote sensing images. Toward this end, we have developed a robust algorithm by combining and localizing feature- and area-based methods. A block-weighted projective (BWP) transformation model is first employed to map the local geometric relationship with weighted feature points in the feature-based stage, for which the weight is determined by an inverse distance weighted (IDW) function. Subsequently, the outlier-insensitive (OIS) model aims to further optimize the registration in the area-based stage. Considering the inevitable outliers (e.g., cloud, noise, land-cover change), OIS integrates Huber estimation with the structure tensor (ST), which is an approach that is robust to residual errors and outliers while preserving edges. Four pairs of remote sensing images with varied terrain features were tested in the experiments. Compared with the-state-of-art algorithms, the proposed algorithm is more effective, in terms of both visual quality and quantitative evaluation.

1. Introduction

Due to the complementary information available in different remote sensing images of the same scene, it is valuable to analyze them jointly for the interpretation and utilization. First of all, it is a vital precondition to ensure geometrical consistency of the ground truth. This is always achieved by registration, which is the process of aligning different images of the same scene acquired at different times, viewing angles, and/or sensors (Wong and Clausi, 2007). As a fundamental preliminary task; registration plays an important role in remote sensing image applications such as image fusion (Shen et al., 2016; Duran et al., 2017), image mosaicing (Li et al., 2015; Jiang et al., 2017), multitemporal image analysis (Gong et al., 2008; Zeng et al., 2013), etc. Over the past decades, quite a few registration algorithms have been proposed and widely applied. These algorithms can be classically divided into two main categories: area-based methods (ABM) and the feature-based methods (FBM) (Zitová and Flusser, 2003).

ABM, which is also regarded as intensity-based method, utilizes the intensity information directly to register images. The emphasis of ABM is the similarity metric (Xu et al., 2016). And the main similarity metrics are as follows. The sequential similarity detection algorithm (SSDA) is quick and easy to implement, whereas it is unstable with complex geometrical deformation, such as rotation and scaling, etc. (Barnea and Silverman, 1972). Mutual information (MI) is also utilized as a similarity measure (Cole-Rhodes et al., 2003; Kern and Pattichis, 2007). MI indicates the best match between a reference image and the sensed image through its maximum value (Cole-Rhodes et al., 2003; Chen et al., 2003). However, MI is time-consuming, although it is robust to radiation differences (Colerhodes et al., 2001). As for normalized cross correlation (NCC), it determines the registration by the relativity of the two images or windows to be matched (Hel-Or et al., 2011). Due to its robust linear tone mapping, NCC is widely utilized for image registration, whereas it can fail to handle nonlinear intensity differences (Hel-Or et al., 2011). In summary, the different similarity metrics are suitable for diverse circumstances, and can obtain a high accuracy, as the entire intensity information is utilized. However, ABM has a weak capability to handle large geometrical deformation.

Differing from ABM, FBM is robust to large geometrical deformation, while also being quick, as it processes images with their significant features rather than all of the intensity information. As we know, point, line, and region are common extracted features (Yu et al., 2008; Sui et al., 2015; Goncalves et al., 2011). Inspired by the most representative

---

* Corresponding authors at: School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China (H. Shen). School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China (X. Li).

E-mail addresses: lixinghua5540@whu.edu.cn (X. Li), shenhf@whu.edu.cn (H. Shen).

https://doi.org/10.1016/j.isprsjprs.2019.03.002
Received 27 November 2018; Received in revised form 3 March 2019; Accepted 3 March 2019

scale-invariant feature transform (SIFT) feature (Lowe, 2004), a variety of improved operators followed, including principal component analysis (PCA-SIFT) (Yan and Sukthankar, 2004), affine SIFT (ASIFT) (Morel and Yu, 2009); uniform robust SIFT (UR-SIFT) (Sedaghat et al., 2011), and scale restrict (SR-SIFT) (Yi et al., 2008). To save time; speeded-up robust feature (SURF) extract the salient points (Bay et al., 2006); and KAZE takes advantage of additive operator splitting (AOS) to extract more valuable feature points (Alcantarilla et al., 2012). And recently, the oriented features from accelerated segment test (FAST) and the rotated binary robust independent elementary features (BRIEF) are employed to extract the robust feature point quickly and robustly (Rosten et al., 2010; Calonder et al., 2012). As for feature lines, they are the representation of object contours (Hui et al., 1995), coastlines (Dongseok et al., 1997), roads (Li et al., 1992), and so on. The mature line segment detector (LSD) has been published online[1] (Gioi et al., 2012) and extensively applied to image alignment and stitching (Xiang et al., 2016; Li et al., 2015; Lin et al., 2017). Furthermore, feature regions can be extracted to design the mapping function according to their gravities (Goshtasby et al., 1986), which are rarely applied for image registration. Comparatively speaking, feature points are more frequently utilized owing to their insensitivity to various distortions, and for the convenience of further calculation. For feature matching, the exhaustive search method (Lowe, 2004) and the KD-tree matching algorithm (Silpa-Anan and Hartley, 2008) are popularly used in practice. The former is a linear search method, which compares the feature point similarities in two images one by one. Thus, it is time-consuming, and it does not take the structural information of features into consideration. Inversely, the KD-tree matching algorithm matches features by designing an index structure, which is an efficient approach, even for large images. With matched features, a transformation model, which can be affine, polynomial, or projective, is employed to rectify the images. FBM can save time, with many speedup strategies; however, owing to outliers, it is difficult to extract features exactly, resulting in a lower accuracy than ABM (Liu et al., 2014).

In terms of the choice between ABM or FBM, the traditional registration algorithms rectify images with a global transformation model (Gong et al., 2014; Liang et al., 2014). However, for wide-field medium and high spatial resolution imagery with complicated terrain, a global relationship cannot correct the inconsistent regional deformation. Under this circumstance, some local transformation models have been proposed in the FBM framework, focusing on the regional deformation, for a higher precision (Goshtasby, 1987; Zaragoza et al., 2014; Liu et al., 2016; Flusser, 1992; Goshtasby, 1986; Nejati et al., 2010; Zagorchev and Goshtasby, 2006). For example, thin plate spline (TPS) takes the regional deformation into consideration by combining a local non-affine and global affine transformation model (Bookstein, 1989; Bentoutou et al., 2005). Since TPS is feature point dependent, outliers and insufficient feature points may result in the topology not being well preserved and a poor result being obtained (Holden, 2008). The piecewise linear model (PLM) estimates the local affine transformation model for every triangle with three feature points (Ye et al., 2017; Han et al., 2014). However, since only three feature points are used for the affine model estimation, the model accuracy cannot be guaranteed. Within each image grid, the As-Projective-As-Possible (APAP) algorithm (Zaragoza et al., 2014) designs local mapping function according to the corresponding Gaussian weight. For digital images in the computer vision field, APAP usually achieves a satisfactory effect. However, for remote sensing images with complicated terrain fluctuations, their individual features should be further considered. Generally speaking, on the one hand, without taking the unique characteristics of medium/high spatial resolution remote sensing images into account, these local transformation models may not be suitable for their registration. On the other hand, all the aforementioned algorithms neglect the outliers such as cloud, noise, and land-cover change, which hinder the registration accuracy, especially for the local transformation model.

In this paper, to cope with the above-mentioned problems, we propose a robust registration algorithm focusing on a local transformation model, while simultaneously considering outliers. The main contributions of this paper include two aspects. One is that we propose the block-weighted projective model on the basis of feature-based method (BWP-FBM), allowing for locally inconsistent deformations in the first stage. The other contribution is that we propose the outlier-insensitive model based on area-based method (OIS-ABM) to optimize the aforementioned result that is subject to inevitable outliers. Specifically, the outlier-insensitive model (OIS) is formed on the basis of Huber estimation and the structure tensor (ST) at the finer scale. For brevity, the combination of BWP-FBM and OIS-ABM makes up the proposed method.

The rest of this paper is organized as follows. Section 2 describes the proposed algorithm in detail. The experimental results are provided in Section 3, where the algorithm verification is also given. Finally, our conclusions are presented in Section 4.

2. Methodology

In this paper, a robust registration algorithm is proposed for the precise local mapping function, as shown in Fig. 1. The emphasis is put on the two stages of BWP-FBM and OIS-ABM. In the BWP-FBM stage, the candidate feature points in yellow are extracted by SIFT and correspondingly matched with each other. Based on the feature points, the regional weighted projective transformation model is estimated in the regularly divided blocks. Since the first stage is equivalent to the coarse registration, the block size shouldn't be too small for time-saving. And the weight of each feature point is assigned by the inverse distance weighting (IDW) function, with the distance from the block center to the point. After resampling, the registration result of BWP-FBM is obtained. Successively, the distinct outliers are primarily eliminated by ST in the OIS-ABM stage. The stage is different from the first one, for high accuracy, the smaller block size should be taken into account. For each smaller block, the Huber estimation is further employed to alleviate the influence of the remaining outliers on the transformation model. By the iterative update process of the transformation model, the optimized result is exported after coordinate transformation and resampling. More details of the algorithm are provided in the following.

2.1. Feature-based method with block-weighted projective model

In the BWP-FBM stage of the proposed algorithm, the potential feature points in the reference image are detected by SIFT and matched with the corresponding points in the sensed image (Lowe, 2004). After candidate feature point matching, random sample consensus (RANSAC) is used to remove the mismatched points and preserves the optimal inner points (Wong and Clausi, 2007).

With the accurate matched point pairs, we focus on weight construction to improve the performance of the local projective model. With the selected weight function, the objective function is described as shown in Eq. (1) and used to estimate the optimal transformation model (Zaragoza et al., 2014; Xiang et al., 2018; Feng et al., 2017):

$$\widetilde{p} = \arg\min_{\widetilde{p}} \sum_{i=1}^{N} \|w_i m_i \widetilde{p}\|^2 = \arg\min_{\widetilde{p}} \|WM\widetilde{p}\|^2 \quad s.\ t.\ \ \|\widetilde{p}\| = 1 \tag{1}$$

where the vector $\widetilde{p}$ represents the geometric relationship between the reference and sensed image. N is the number of matched point pairs, w_i is the weight of the $i - th$ matched feature point, and m_i is the exterior product of the corresponding feature points. $W \in \mathbb{R}^{2N\times 2N}$ is a diagonal matrix, of which the main diagonal element is w_i, and $M \in \mathbb{R}^{2N\times 9}$ is obtained by stacking m_i vertically for all of i.

In Eq. (1), the weight function is introduced to balance the

---

[1] http://demo.ipol.im/demo/gjmr_line_segment_detector/.

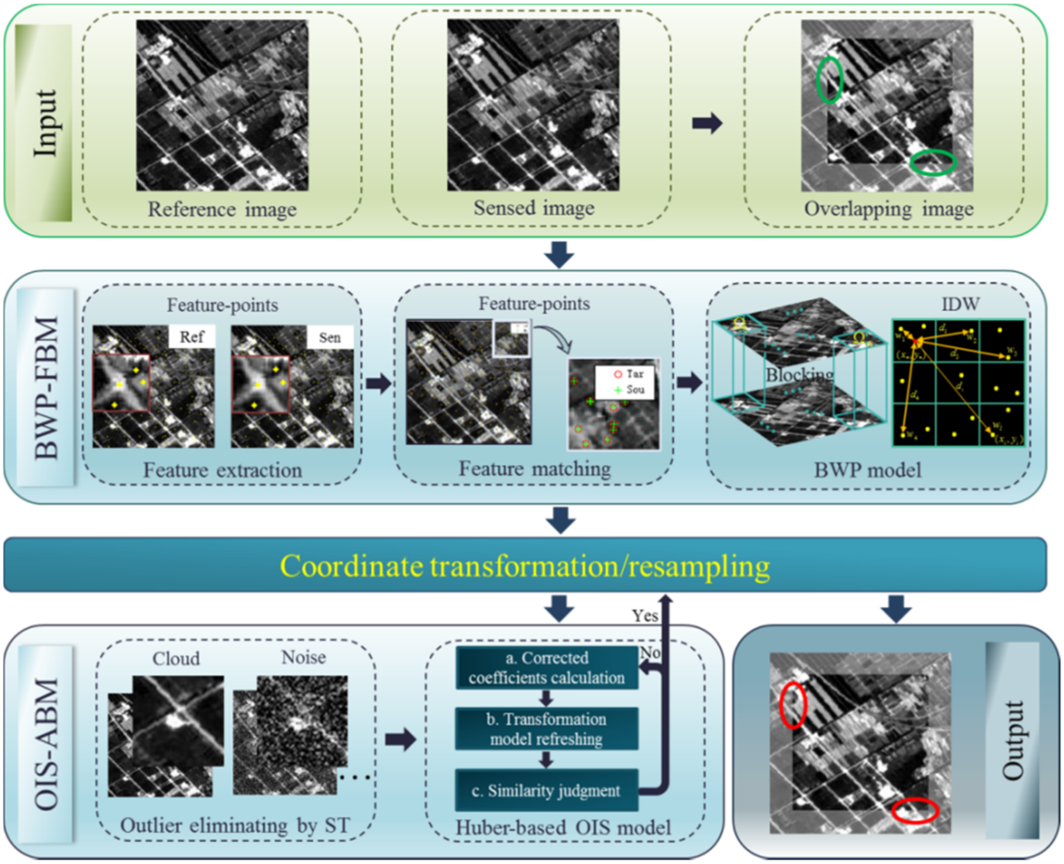


Fig. 1. Flowchart of the proposed method for remote sensing image registration.

contribution of the points in and around the block, as they all affect the accuracy of the transformation model. This is inspired by the work of Zaragoza et al. (2014) in digital image stitching. As mentioned in their work, the Gaussian function is employed to distribute weight to different feature points and the upper-left corner of cell is the reference point. However, in our work, the block center is set as the reference point and the IDW function is employed to distribute the weights to different feature points. As known, IDW is one of the most commonly used deterministic models in spatial interpolation based on the first law of geography (Tobler, 1970; Lu and Wong, 2008). The weight is calculated as follows:

$$w_i = \frac{1/\sqrt{(x_i - x_*)^2 + (y_i - y_*)^2}}{\sum_{i=1}^{N} 1/\sqrt{(x_i - x_*)^2 + (y_i - y_*)^2}} \tag{2}$$

where (x_i, y_i) is the $i - th$ feature point coordinate in the reference image. N is also the number of matched feature point pairs. (x_*, y_*) is the block center coordinate.

Fig. 2 shows the weight assignment of the feature points, using nine blocks as an example. For the first block marked by a red rectangle at the top left corner, the red dot represents the block center, and the black dots are the simulated feature points. When calculating the transformation model, the distance from the red dot to the black dot is calculated as $d_i (i = 1, 2, 3, \cdots)$. The weight $w_i (i = 1, 2, 3, \cdots)$ is then determined by IDW.

For brevity, all the Cartesian coordinates are transformed to homogeneous coordinates. Basically, direct linear transformation (DLT), which is similar to the least-squares algorithm, estimates the projective transformation model from a set of matched feature points (Xiang et al., 2018; Hartley and Zisserman, 2003). If the transformed point in the reference image is coincident with the original corresponding point in the sensed image, the corresponding relationship can be rewritten by exterior product as $X' \times PX = \vec{0}_{3\times1}$ and linearized as shown in Eq. (3):

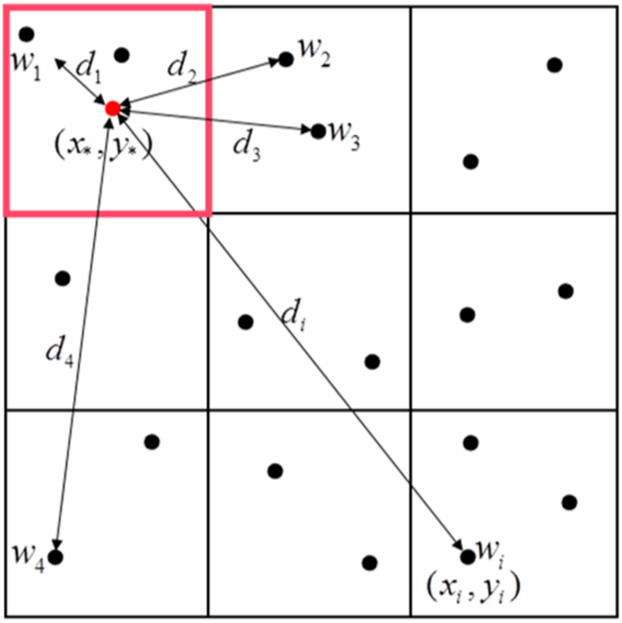


Fig. 2. Weight distribution for the BWP model.

$$\begin{bmatrix} \vec{0}_{1\times3} & -X^T & y'X^T \\ X^T & \vec{0}_{1\times3} & -x'X^T \\ -y'X^T & x'X^T & \vec{0}_{1\times3} \end{bmatrix} \begin{bmatrix} p_1^T \\ p_2^T \\ p_3^T \end{bmatrix} = \vec{0}_{3\times1} \tag{3}$$

where $P = [p_1\ p_2\ p_3]^T$ is the projective transformation model, which is converted to a column vector denoted as $\widetilde{p}$ in Eq. (1). For every matched feature point, Eq. (3) is constructed, and the leftmost matrix is denoted as m_i in Eq. (1).

Nevertheless, Eq. (1) may be unstable when the feature points distribute sparsely in the block and the surrounding feature points are far away from the block center. To deal with this situation, we compensate the weights with a small constant $w_{off} \in [0, 1]$. In this way, the corrected weight w_i^{corr} of each point is obtained by Eq. (4):

$$w_i^{corr} = \max(w_i, w_{off}) \tag{4}$$

The equation is applied to each pair of matched points, and the solution to Eq. (1) is the least significant right singular vector of WM. With the precise BWP model, the coordinates are transformed and resampled by bilinear interpolation. Finally, the registration result is obtained.

2.2. Area-based method with outlier-insensitive model

The aforementioned result is optimized by the OIS model. Due to cloud, noise, and land-cover changes, Huber estimation is employed to design the mapping function, with the inner pixels filtered by ST iteratively at a finer scale.

Huber estimator is one of the popular M-estimation functions (Li et al., 2014; Shen and Zhang, 2009). If we suppose that the transformation model coefficients to be solved are denoted as $\theta = [\theta_1, \theta_2, ..., \theta_i]^T$, then the M-estimation is described as shown in Eq. (5):

$$\hat{\theta} = \arg\min\left(\sum_i \rho(r_i, \theta)\right) \tag{5}$$

where r_i is the residual error of the $i-th$ pixel, and ρ is the objective function.

With the Huber estimator, ρ is defined as shown in Eq. (6). The piecewise objective function is constructed according to the threshold t_1 and the intensity residual error r_i. When it is smaller than the threshold, the quadratic function is utilized to estimate the corrected coefficients [see Eq. (6)]. Otherwise, the linear function is employed, which aims to get rid of the influence of the ignored outliers [see Eq. (6)].

$$\rho(r_i) = \begin{cases} r_i^2/2 & \text{while } |r_i| < t_1 \\ t_1(|r_i| - t_1/2) & \text{while } |r_i| \geqslant t_1 \end{cases} \tag{6}$$

where $t_1 = 1.345 \times \sigma$, and σ is the standard deviation of all r_i. Fig. 3 shows the variation curves with different thresholds whose X-axis indicates the intensity residual errors and Y-axis denotes the Huber loss.

Generally speaking, the minimization of Eq. (5) is converted to iteratively reweighted least-squares (IRWLS) process, and the derivative of the objective function is utilized to calculate the weight function. The first derivative of objective function $\rho(r_i)$ is referred to as an influence function $\psi(r_i)$, as shown in Eq. (7):

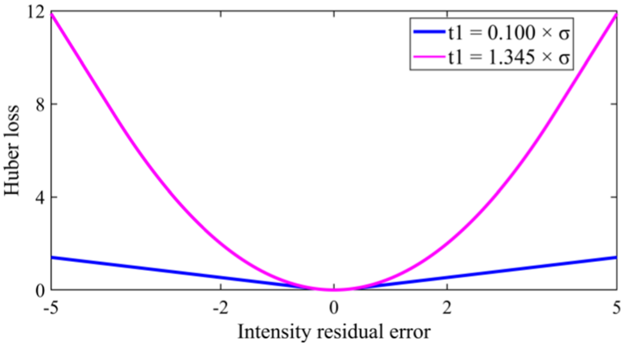


Fig. 3. The Huber estimator with different thresholds.

$$\psi(r_i) = \begin{cases} r_i & \text{while } |r_i| < t_1 \\ t_1 \operatorname{sgn}(r_i) & \text{while } |r_i| \geqslant t_1 \end{cases} \tag{7}$$

where ‘sgn’ means that it is a sign function.

By dividing the influence function $\psi(r_i)$ by the residual error r_i, the weight function is obtained as shown in Eq. (8):

$$w(r_i) = \psi(r_i)/r_i = \begin{cases} 1 & |r_i| < t_1 \\ t_1/|r_i| & |r_i| \geqslant t_1 \end{cases} \tag{8}$$

As seen, the pixel will be endowed with a small weight if the residual error is large. Otherwise, the weight is equal to one. Since the Huber estimation is converted to IRWLS process, the optimal corrected coefficients are calculated as shown in Eq. (9):

$$Y^k = \arg\min\left(\sum_{i=1}^{h*l} w_i^k r_i^k r_i^k\right) \tag{9}$$

where Y^k is the corrected coefficients of the transformation model in the $k-th$ iterative process, where representing $(dh_0^k, dh_1^k, da_0^k, ..., db_2^k, db_3^k)$ as the linear polynomial model is employed to describe the geometric relationship of the two images. The algorithm is performed in an assigned window with the size of $h-by-l$ pixels.

As for the aforementioned residual error, it points to the intensity residual error of the corresponding pixel (Afsharnia et al., 2017), which can be described as shown in Eq. (10):

$$r_i = h_0 + h_1 I_2(a_0 x_i + a_1 y_i + a_2 x_i y_i + a_3, b_0 x_i + b_1 y_i + b_2 x_i y_i + b_3) - I_1(x_i, y_i) \tag{10}$$

where $(a_0, a_1, a_2, a_3, b_0, b_1, b_2, b_3)$ is the parameter of the linear polynomial model. (x_i, y_i) are the coordinates in the sensed image. I_1 and I_2 represent the reference and sensed image’s intensity, respectively. h_0 and h_1 are linear radiation coefficients. The initial coefficients of Eq. (10) were set as [0, 1, 1, 0, 0, 0, 0, 1, 0, 0] in our experiments.

As seen, Eq. (9) is a nonlinear function, which can be solved by Taylor expansion. The iterative formula is described as shown in Eq. (11):

$$\theta^{(k+1)} = \theta^k + Y^{(k+1)} \tag{11}$$

where θ^k and $\theta^{(k+1)}$ are linear polynomial coefficients calculated in the $k-th$ and $(k+1)-th$ iteration, respectively.

Specifically, the correlation coefficient (CC) is used as the termination condition in the iteration. If the CC in the$(k+1)-th$ iteration is smaller than that in the $k-th$ iteration, the transformation model in the $k-th$ iteration is the best. Otherwise, the iteration continues. Until the maximal CC is obtained, the linear polynomial transformation model is the most optimal for image rectification.

Despite restraining the large residual error with a small weight in the IRWLS process, Eq. (9) may be unstable when the pixel weight is too small. Since most of these pixels are obvious outliers, such as cloud, noise, and so on, they can be detected and removed with ST before estimating the OIS model by Huber estimation.

By taking the geometrical structure into consideration, ST is widely used to detect the edges and corners of image (Köthe, 2003). Here, it is utilized to detect outliers while preventing the edges from being regarded as outliers. The algorithm can be described as shown in Eq. (12):

$$\begin{cases} |r_i| < t_2 \times st_{I_1(x,y)} & (x, y) \text{ is inner} \\ |r_i| \geqslant t_2 \times st_{I_1(x,y)} & (x, y) \text{ is outlier} \end{cases} \tag{12}$$

where t_2 is a constant with a large range; $st_{I_1(x,y)}$ is ST at point (x, y) in image I_1; and “inner” means that the pixel will be brought into calculation, or otherwise rejected as “outlier”.

In addition, ST is always calculated with the gradient in the selected window. Firstly, the gradient and the Gaussian average-gradient covariance matrix are calculated as shown in Eq.(13):

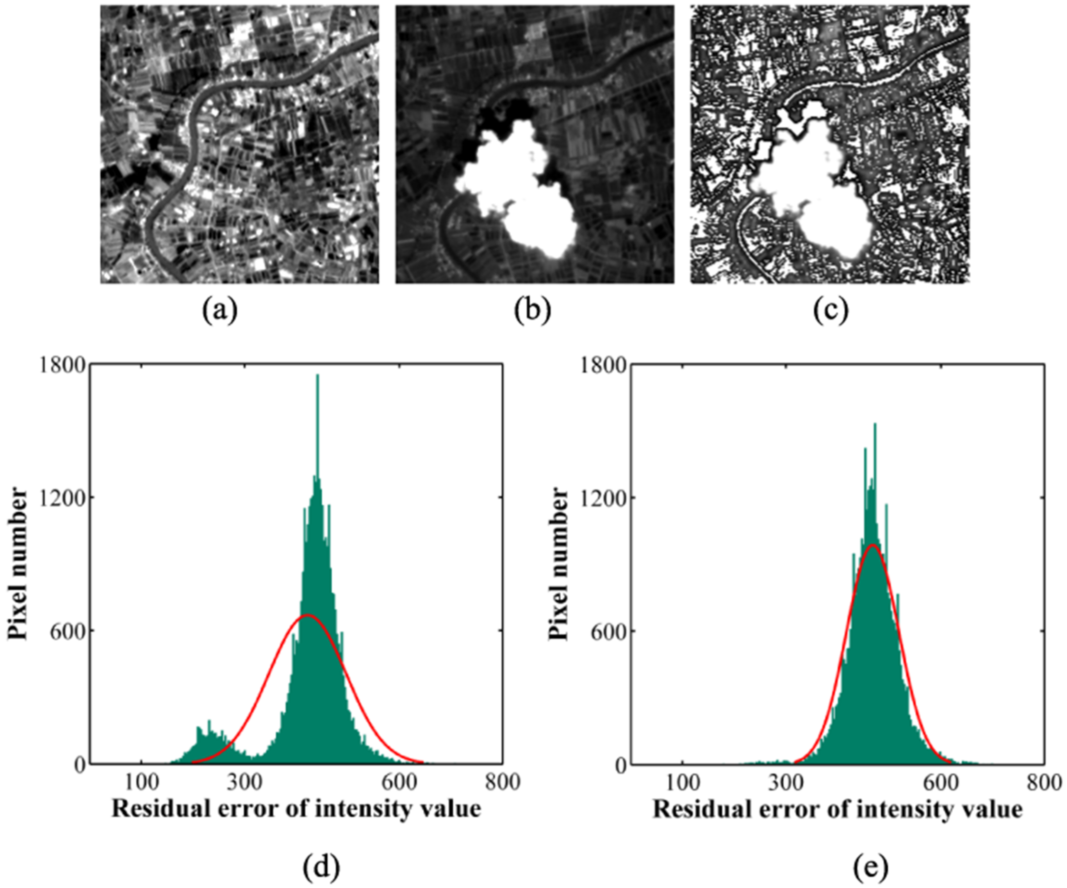


Fig. 4. ST (a) Reference image. (b) Sensed image. (c) Outliers. (d) Residual error distribution before outlier elimination. (e) After outlier elimination.

$$G = \begin{bmatrix} G_x \\ G_y \end{bmatrix}$$

$$\bar{C} = G_\sigma * \begin{bmatrix} G_x^2 & G_x G_y \\ G_y G_x & G_y^2 \end{bmatrix} \tag{13}$$

where G denotes the gradient; and G_x and G_y are the gradients in the x and y directions, respectively. $\bar{C}$ is the Gaussian average-gradient covariance matrix, and $G_\sigma$denotes the Gaussian filtering with standard deviation σ.

ST is then the sum of the absolute eigenvalues calculated by singular value decomposition, and is denoted as "st", where:

$$st = |\lambda_1| + |\lambda_2| \tag{14}$$

where λ_1 and λ_2 are the eigenvalues of $\bar{C}$.

In order to illustrate the effectiveness of ST, a pair of images is employed [see Fig. 4(a) and (b)]. In Fig. 4(b), there are radiation differences, geometrical deformation, cloud, and land-cover changes. As can be seen, the residual error of the two images is not a normal distribution [see Fig. 4(d)] before ST detects outliers. Registering the two images by Huber estimation directly, Eq. (9) may be unstable and cause a large error. After ST employed, the outliers are clearly marked in Fig. 4(c) in white. After the outlier removal with the constant of 20, the residual error obeys a normal distribution in Fig. 4(e).

When filtering all the pixels with ST, most of the outliers are rejected and the residual error obeys a normal distribution. Generally speaking, ST is integrated with the Huber estimation, in which way all the pixels can be given different weights to construct the OIS model according to Eqs. (8) and (12). The weight of each iteration is calculated as follows:

$$w_i = \begin{cases} 0 & |r_i| \geqslant t_2 \times st_i \\ t_1/|r_i| & t_1 \leqslant |r_i| \leqslant t_2 \times st_i \\ 1 & |r_i| < t_1 \end{cases} \tag{15}$$

where i denotes the $i-th$ pixel in the image.

Once the optimal transformation model is acquired, the coordinates are transformed and bilinear interpolation is employed to calculate the gray values of the corresponding pixels. After successful coordinate transformation and resampling, the optimized registration result is obtained.

3. Experiments and evaluation

We compared the proposed method with the following five registration algorithms, from the perspectives of both the qualitative and quantitative results. A representative of the conventional registration algorithms is the global projective transformation model (GPTM) (Ye and Shan, 2014). PLM, which is integrated into ENVI, divides the image into triangles and estimates the affine transformation model for each triangle, (Goshtasby, 1986; Ye, 2017). The APAP algorithm in (Zaragoza et al., 2014) constructs a projective transformation model with blocks. The TPS model calculates the deformation per pixel with matched feature points, which can be performed in ArcMap. The outcomes of BWP-FBM and the proposed method were also compared. For the qualitative evaluation, i.e., visual quality, on the one hand, overlapping the reference and aligned image allowed us to conveniently observe the overall performance. On the other hand, observing the enlarged images with view swipe can allow us to evaluate the alignment in detail. Notably, the reference image appears in true color, whereas the sensed and aligned images appear in pseudo color (the NIR, red, and green bands are assigned to the red, green, and blue channels,

Table 1
Overview of the experimental data.

Label	Res	Time	Sensor	Size	Block	Characteristics
Mountain	16 m	2017.07.08 2017.09.20	GF-1_{WFV-2}	400 × 400 × 4	200 × 200 pixels 12 × 12 pixels	Complex terrain features to the northwest of Beijing, China.
Urban area	30 m	1999.08.11 1997.08.13	ETM+ TM	1000 × 1000 × 6	200 × 200 pixels 50×50 pixels	Dense buildings surrounded by farmland in Tianjin, China.
Farmland	16 m	2016.05.12 2015.08.07	GF-1_{WFV-2}	1000 × 1000 × 4	200 × 200 pixels 25 × 25 pixels	Homogeneous texture, lakes, and clouds in Wuhan, China
Hybrid area	3.7 m	2017.10.09 2016.10.11	TRIPLESAT-2_{MS} GF-2_{PMS1}	1084 × 1084 × 4	216 × 216 pixels 50 × 50 pixels	Vegetation, bare area, and buildings in Nanning, China

Note: "Res" represents the spatial resolution.

respectively) (Han et al., 2015). Furthermore, the yellow arrows and points are labeled in the enlarged sub-regions to allow convenient observation of the deformation. The direction of the arrow indicates the dislocation of the aligned image relative to the reference image, and the length is a rough quantitative representation of the deformation. A yellow point suggests that the processed image is well aligned to the reference image. In addition, the results were quantitatively evaluated by the three indicators of CC, normalized mutual information (NMI), and root-mean-square error (RMSE). As far as each perspective was concerned, four realistic experiments with different terrain features were conducted.

Details of the experimental images are provided in Table 1, including the spatial resolution ("Res"), imaging time, size etc. For brevity, each experimental image pair is marked with their own main feature, such as "Mountain", "Farmland", or "Urban area". The block sizes in the two stages of the proposed method are also listed. Further details are provided as follows.

3.1. Visual quality

The four experiments were conducted and evaluated by the visual quality, to check the alignment result. Before focusing on the registration results, there's an illustration of "overcompensation" that is a frequently used word in our experimental description. It is a phenomenon in the registration process. Taking Fig. 5 for example, the road (denoted by "r_2") in the sensed image [Fig. 5(b)] has a two-pixel deviation [see Fig. 5(c)] to its counterpart (denoted by "r_1") in the reference image [Fig. 5(a)]. If we shifted the sensed image by two pixels to the right, and the accurate registration result could be obtained, like Fig. 5(d). When it is moved right by more than two pixels, this causes the overcompensation [see Fig. 5(e)]. That's to say, overcompensation means that the movement direction for accurate registration is correct but the displacement is larger than the real value.

As seen, the first experiment was conducted with the "Mountain" image pair, as listed in Table 1. As shown in Fig. 6, the image contains abundant ridges and valleys. Fig. 6(c) is the overlay of the reference image [Fig. 6(a)] and sensed image [Fig. 6(b)]. This image is fuzzy and there is much ghosting, which is caused by the misalignment. However, Fig. 6(d), which is made up of the reference image and the aligned result by the proposed method with 396 pairs of matched feature points, is visually clear and the ghosting is mostly eliminated. To further evaluate the registration of the different algorithms, the yellow rectangle regions from the view swipe of the reference image and the results of GPTM, TPS, PLM, APAP, and the proposed algorithm are enlarged, respectively. Simultaneously, the similar regions of the original images are extracted and enlarged in Fig. 6(e). For region I in Fig. 6(e), the horizontal offset between the reference and sensed image is about two pixels. In addition, a vertical offset of one pixel is apparent in region II of Fig. 6(e). Firstly, in region I, the horizontal offset is not eliminated by GPTM, as shown in Fig. 6(f). However, overcompensation is achieved by TPS in Fig. 6(g), causing dislocation while enlarging the marked region I. There is still a one-pixel horizontal offset between the reference and aligned image by PLM in Fig. 6(h). The offset is not corrected by APAP in Fig. 6(i), whereas it is eliminated by the proposed method in Fig. 6(j). Secondly, in region II, the result of the proposed method in Fig. 6(j) outperforms that of the other methods. For the result of APAP in Fig. 6(i), overcompensation is apparent. However, PLM almost completely eliminates the vertical dislocation in Fig. 6(h). Fig. 6(g) shows that there is a less than one-pixel offset between the reference and aligned image. There is also a one-pixel offset in the vertical direction in Fig. 6(f) and (i), which are processed by GPTM and APAP, respectively. Both horizontally and vertically, GPTM [Fig. 6(f)] performs worse than the other methods. This is because, in a mountainous area, the features always distribute sparsely and unevenly, and the global transformation model cannot describe the local deformation correctly. Overall, the partial model performs well allow for the local inconsistent deformation of mountain area, specially the proposed method can eliminate most of the inconsistent distortions.

The "Urban area" image pair was used in the second experiment. The experimental area contains buildings surrounded by agricultural area [see Fig. 7(a) and (b)]. The experimental data are from ETM+ and TM images taken at different times. There are not only geometrical deformations, but also land-cover changes apparent in Fig. 7(c). Processed by the proposed method with 374 pairs of matched feature points, the overlapping regions are almost aligned with each other [see Fig. 7(d)]. The five registration algorithms were again conducted. In particular, Fig. 7(e) comes from the checkerboard of original reference and sensed image. They cover a much wider area than the other enlarged images in Fig. 7(f)–(j) as there is large deformation and corresponding contents to Fig. 6(f)–(j) are marked in white dotted rectangles in Fig. 7(e). TPS can eliminate the vertical offset shown in Fig. 7(g) although the edge of the road is not clear, which may be a result of the resampling method after coordinate transformation. All the algorithms can eliminate the large deformation between the reference and sensed image. However, there are some small horizontal displacements left in results processed by GPTM in Fig. 7(f) and PLM in Fig. 7(h). The horizontal overcompensation is apparent in the results of TPS in Fig. 7(g). Although APAP aligns the road in the sensed image to the reference image, the road breaks under the edge of the checkerboard in Fig. 7(i). In region II, the enlarged images show us the vertical alignment. The road from top to bottom is disconnected in Fig. 7(f) and (h) processed by GPTM and PLM, respectively. Overcompensation is again apparent in the result obtained by APAP, as shown in Fig. 7(i). As expected, the proposed method gives a desirable outcome in Fig. 7(j). The proposed method divides the image into blocks according to the terrain features. With all the feature points employed, varied weights are assigned to the different feature points based on the contribution to the transformation model. Furthermore, the registration result is further refined in the second stage. Finally, the optimal outcome is acquired by the proposed method.

In the first two experiments, the proposed method performed well on the images without distinct outliers. To further verify the

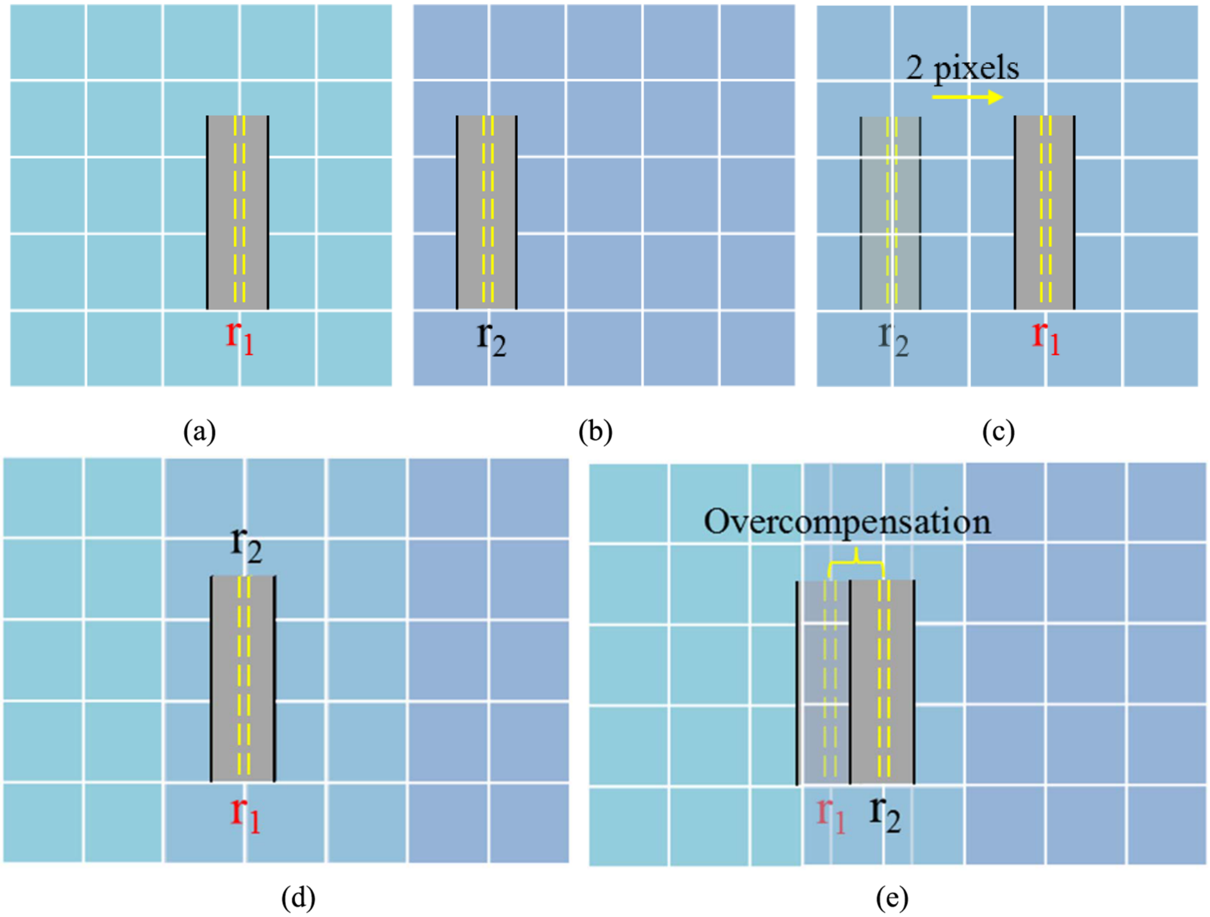


Fig. 5. Illustration of the overcompensation. (a) The reference image. (b) The sensed image. (c) The overlap of the original images. (d) The accurate registration result. (e) The overcompensation result.

effectiveness of the proposed method, the "Farmland" image pair containing cloud, noise, and land-cover changes was employed here. Fig. 8(a) is the reference image and Fig. 8(b) is the sensed image. When overlapping the two images in Fig. 8(c), the result is slightly fuzzy as there is some small deformation between the reference and the sensed image. It is not obvious whether the deformation is thoroughly eliminated in Fig. 8(d) processed by the proposed method with 54 pairs of matched feature points. So we focus on the enlarged regions in Fig. 8(e)–(j), which represent the regions marked by the yellow rectangles in Fig. 8(c) processed by the different algorithms. The extracted sub-regions of the original images are enlarged in Fig. 8(e), where the vertical and horizontal offsets are about two pixels between the reference and sensed image. When focusing on region I, the horizontal offset is not eliminated by GPTM [Fig. 8(f)], and it is the same as the original images in Fig. 8(e). Larger offsets are offered by TPS in Fig. 8(g) and PLM in Fig. 8(h). It is apparent that APAP cannot eliminate the horizontal offset in Fig. 8(i) as well. For region II, the road goes through it, and the emphasis is put on the road edge. The vertical deformation has been eliminated by GPTM in Fig. 8(f) and the proposed method in Fig. 8(j). However, TPS is not able to eliminate the vertical offset between the reference and sensed image [Fig. 8(g)], which is the same as the original images in Fig. 8(e). The original sensed image was moved down by PLM in Fig. 8(h) whereas there is still a small translation to be compensated for. APAP results in overcompensation, leading to vertical deformation in the opposite direction. Based on SIFT, it is difficult to extract enough feature points from the repeated texture area of "Farmland". Thus, the feature points distribute unevenly, especially on the lakes and cloudy areas. In this case, TPS and PLM give a poor performance. GPTM obtains a desirable result in some regions, and not in others. However, the proposed algorithm outperforms all the other methods. The reason for this is that all the feature points are employed and varied weights are assigned to the different feature points based on the contribution to the transformation model. Therefore, the geometric relationship is more accurate and a precise result is obtained.

Actually, a mixed-terrain area is more common than a single terrain type. Among the five methods, which is effective for registering complex terrain images? To answer this question, a pair of images consisting of vegetation, buildings, lakes, and bare land was employed, where there are obvious land-cover changes as well, namely, the "Hybrid area" image pair. Fig. 9(c) is the overlay image of the reference [Fig. 9(a)] and sensed image [Fig. 9(b)]. There is obvious deformation according to the heavy ghosting in Fig. 9(c) whereas it eliminated in Fig. 9(d) processed by the proposed algorithm with 532 pairs of matched feature points. As previously, the yellow rectangle regions marked in Fig. 9(c) from the different algorithms results are extracted and enlarged in Fig. 9(e)-(j). The large deformations are apparent between the original reference and sensed image in Fig. 9(e). The two sub-regions of the GPTM result are enlarged in Fig. 9(f). The dislocated road in region II means that this could not be compensated by GPTM. Furthermore, in region I, there is still a one- or two-pixel horizontal offset. TPS can align the images well in some regions, although the result in the other regions is poor. In Fig. 9(g) the road is visually continuous, whereas there is still a one- or two-pixel vertical offset in region II. In Fig. 9(h), the PLM alignment result is reversed. The road in region II is well registered, whereas overcompensation is apparent in region I. An undesirable outcome is obtained by APAP in Fig. 9(i), where most of the horizontal deformation is eliminated, but the vertical displacement is not completely corrected. In addition, the roads are discontinuous and

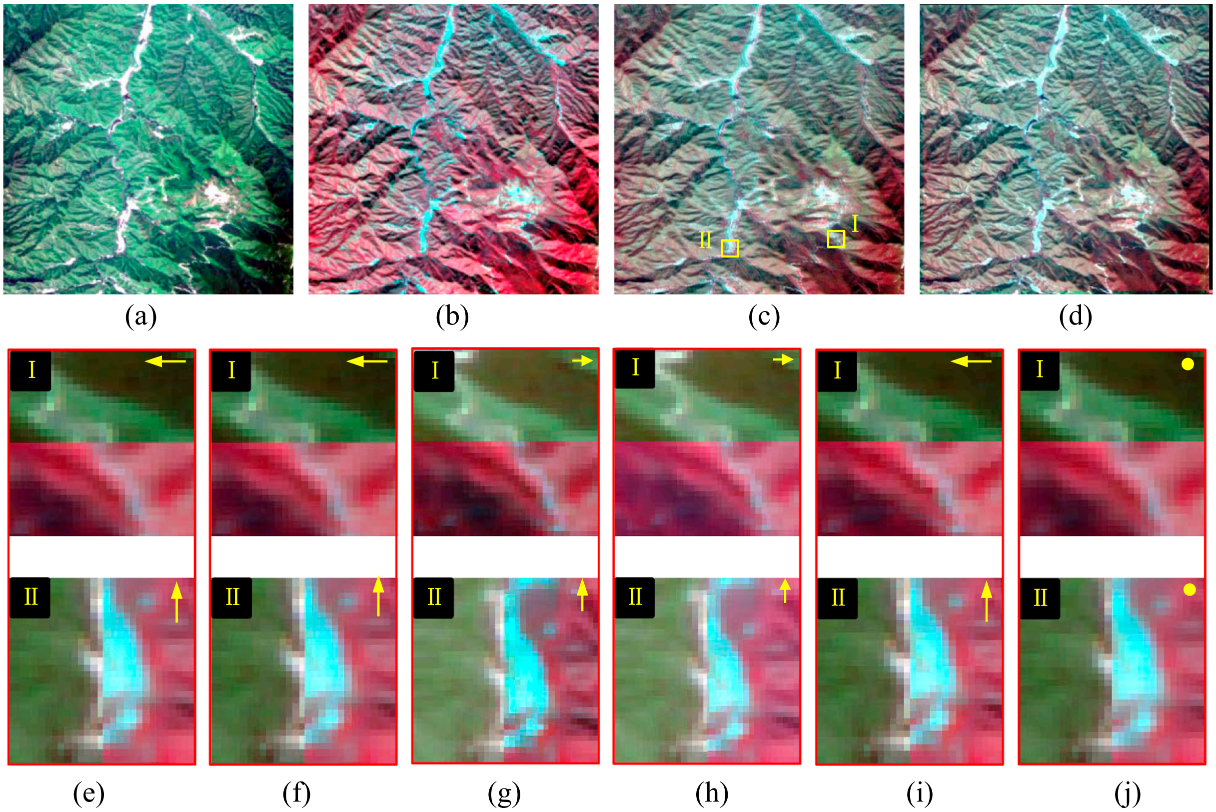


Fig. 6. Registration result for the “Mountain” image pair. (a) Reference image (shown in true color). (b) Sensed image (shown in pseudo color). (c) The overlap of the original images. (d) The overlap of the reference image and the result of the proposed algorithm. Magnified images from (e) the original images, (f) GPTM, (g) TPS (h) PLM, (i) APAP, and (j) the proposed method.

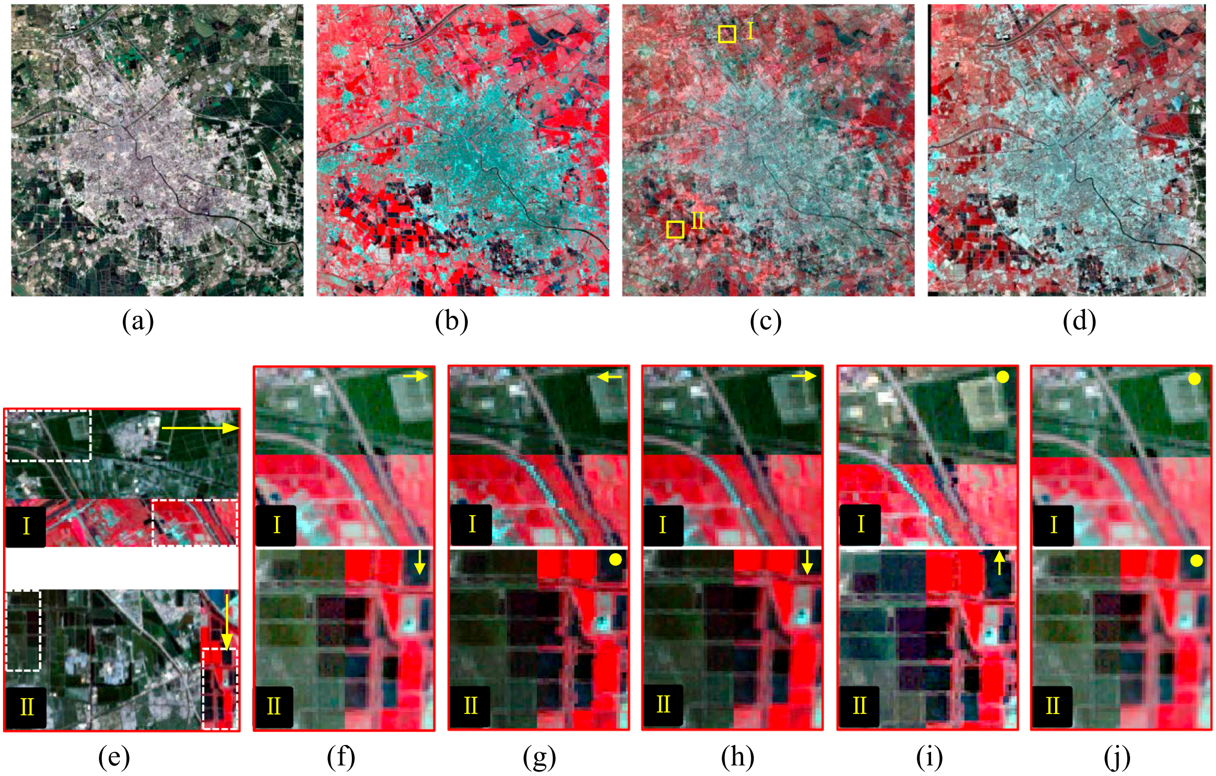


Fig. 7. Registration result for the “Urban area” image pair. (a) Reference image (shown in true color). (b) Sensed image (shown in pseudo color). (c) The overlap of the original images. (d) The overlap of the reference image and the result of the proposed algorithm. Magnified images from (e) the original images, (f) GPTM, (g) TPS (h) PLM, (i) APAP, and (j) the proposed method.

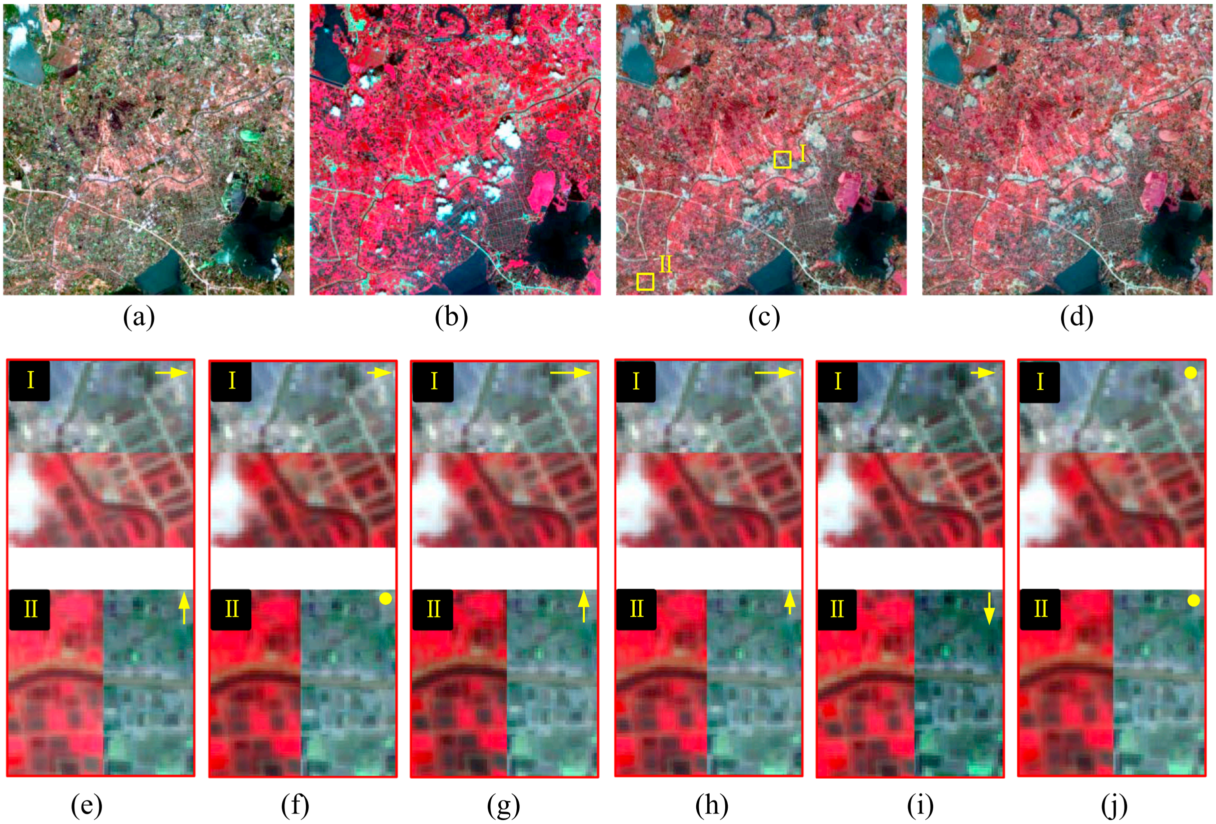


Fig. 8. Registration result for the "Farmland" image pair. (a) Reference image (shown in true color). (b) Sensed image (shown in pseudo color). (c) The overlap of the original images. (d) The overlap of the reference image and the result of the proposed algorithm. Magnified images from (e) the original images, (f) GPTM, (g) TPS (h) PLM, (i) APAP, and (j) the proposed method.

unsmooth in the aligned image, which may be caused by the nearest neighbor interpolation. As expected, a desirable result is obtained by the proposed method in Fig. 9(j), where the roads are well aligned, both horizontally and vertically. Since the spatial resolution is high, the local distortion is more obvious. The proposed method constructs the regional transformation model, allowing for the feature point position errors. Overall, this experiment demonstrates that the proposed method obtains a very satisfactory outcome in this mixed-terrain area.

From the results of the above four experiments, it is apparent that the proposed method is able to align images in both single-terrain area and a mixed-terrain area. To further testify whether the second step (OIS-ABM) is effective for the first step (BWP-FBM) of the proposed method or not, their magnified images in the previous experiments are shown in Fig. 10. The registration accuracy is very approximate between the results of BWP-FBM and OIS-ABM in some area, such as Fig. 10(a) and (c). However, in other areas, a different phenomenon occurs that the obvious improvement provided by OIS-ABM. As seen, in the second row of Fig. 10(b) and (d), OIS-ABM aligns the roads accurately whereas they are misaligned in the first row by BWP-FBM. Therefore, the OIS-ABM further guarantees the high registration accuracy on the whole. In other words, its improvement is relatively effective unless the registration accuracy of the first step is very precise. To further verify the experimental conclusion objectively, the quantitative evaluation is provided in the next section.

3.2. Quantitative evaluation

To give an objective evaluation, the proposed method was compared with the same five methods quantitatively, i.e., GPTM, TPS, PLM, APAP, and BWP-FBM (the first stage of the proposed method). The three assessment criteria of NMI, CC, and RMSE were employed.

(1) NMI. As one of the similarity metrics in ABM, NMI was employed to estimate the overall alignment (Han et al., 2015):

$$NMI(A, B) = \frac{H(A) + H(B)}{H(A, B)} \tag{16}$$

where $H(A)$ and $H(B)$ are the entropies of image A and image B, respectively; and $H(A, B)$ is the joint entropy. The available range of NMI is $[1, 2]$. A larger value of NMI represents a more precise registration result.

(2) CC. The CC metric was used to provide an overall judgement as to whether the sensed image was well aligned to the reference image or not (Han et al., 2015):

$$CC = \frac{\sum_{i=1}^{T}(A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^{T}(A_i - \bar{A})^2}\sqrt{\sum_{i=1}^{T}(B_i - \bar{B})^2}} \tag{17}$$

where A_i and B_i are the intensity values of the $i - th$ pixel in image A and image B, respectively; and $\bar{A}$ and $\bar{B}$ are the corresponding average intensity values. T is the pixel number. The available range of CC is $[-1, 1]$, and CC generally belongs to $[0, 1]$. A larger value of CC means a more accurate registration result.

(3) RMSE. This indicator focuses on evaluating the registration result by calculating the average distance of the corresponding points in the reference and aligned image (Wong and Clausi, 2007; Han et al., 2015; Yang et al., 2017):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta x_i)^2 + (\Delta y_i)^2} \tag{18}$$

where N is the number of evaluated points, and Δx_i and Δy_i are the residual differences of the $i - th$ checkpoint pair in the x and y directions. A smaller RMSE means a better result. For RMSE calculation, a number of evaluated points were extracted from the reference and

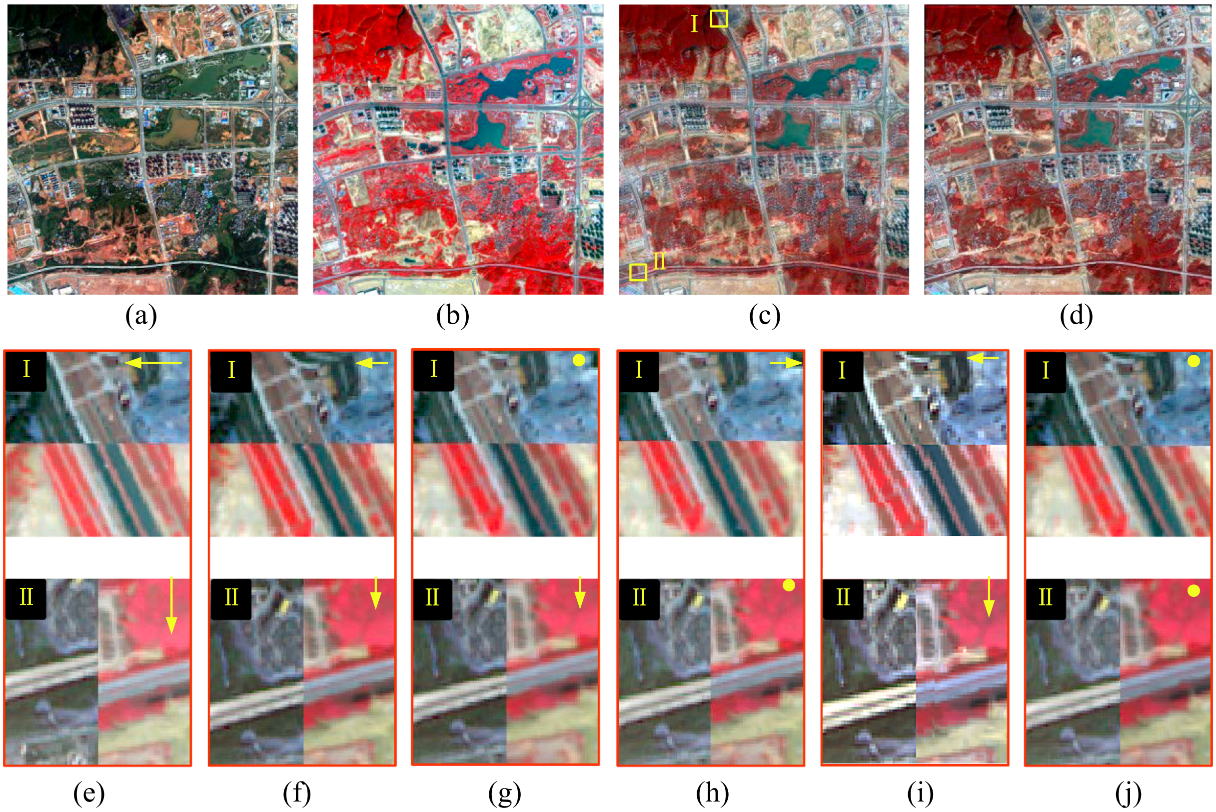


Fig. 9. Registration result for the "Hybrid" image pair. (a) Reference image (shown in true color). (b) Sensed image (shown in pseudo color). (c) The overlap of the original images. (d) The overlap of the reference image and the result of the proposed algorithm. Magnified images from (e) the original images, (f) GPTM, (g) TPS (h) PLM, (i) APAP, and (j) the proposed method.

aligned images humanly.

The CC and NMI evaluation results are shown in Fig. 11. The four lines in different colors represent the evaluations of the four experiments. In Fig. 11(a) and (b), the NMI and CC in the "Farmland" area are lower than those in the other areas. As can be seen, there are distinct clouds in the sensed image, but not in the reference image [see Fig. 8(a) and (b)], and the experimental data were collected at different times, whereas crops are sensitive to temporal changes. From the perspective of locality, both the NMI and CC of TPS in the "Mountain" and "Hybrid area" are lower than those of the other methods. Due to the sparse feature points in mountainous areas, TPS performs poorly in the area with insufficient feature points. In the "Hybrid area", there are distinct land-cover changes, causing fewer matched point pairs. Therefore, TPS cannot obtain satisfactory results in the "Mountain" and "Hybrid" areas. The proposed method obtains the maximum NMI and CC values. In addition, while BWP-FBM achieves a high accuracy, the proposed method shows only a small accuracy difference with BWP-FBM, as in the experiment in the "Farmland" area in Fig. 11. In contrast, there exists a significant improvement from the BWP-FBM method to the proposed method when the former accuracy is not as high as possible, as in the "Mountain" and "Hybrid" experiments in Fig. 11. Although the tendency of the two evaluation indicators is the same, on the whole, there are minor inconsistencies between NMI and CC in the "Farmland" area, when focusing on the TPS result. The two indicators may be

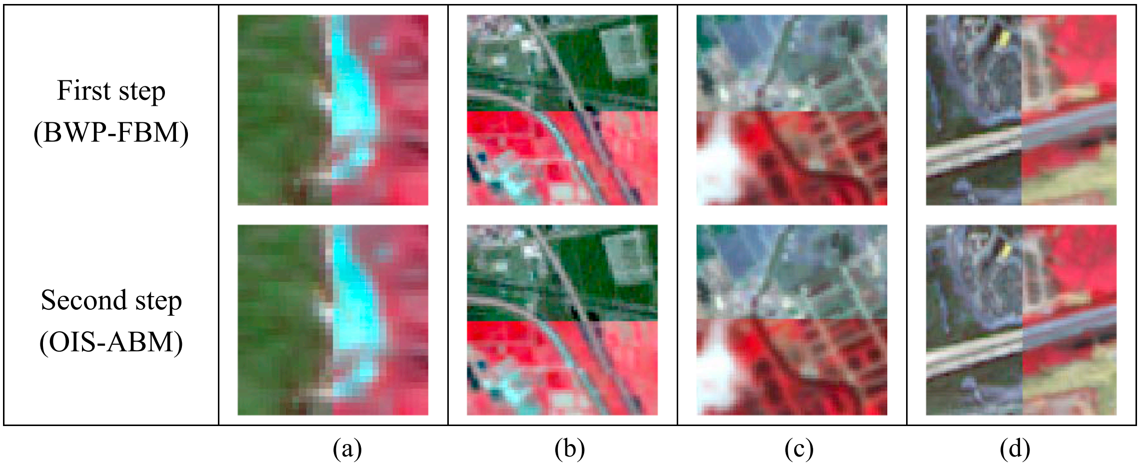
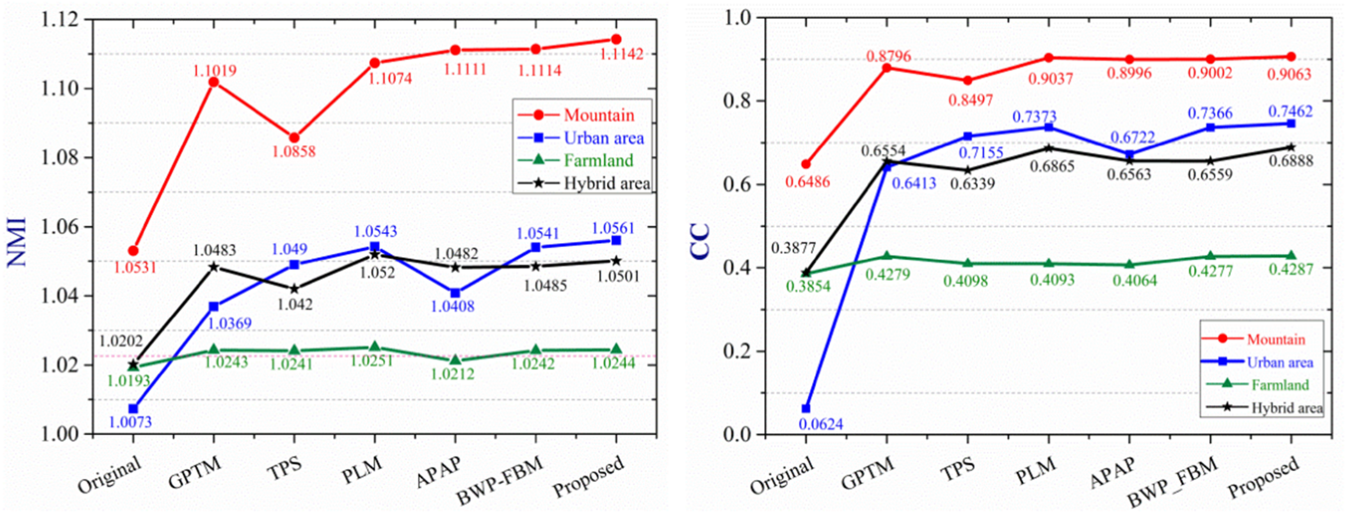


Fig. 10. The comparison of BWP-FBM (in the first row) and OIS-ABM (in the second row) results from the aforementioned experiments. (a) Fig. 6 in region II, (b) Fig. 7 in region I, (c) Fig. 8 in region I, (d) Fig. 9 in region II.



(a)

(b)

Fig. 11. Quantitative evaluation of the registration results in the four experiments. (a) NMI. (b) CC. ("Original" means evaluation on the unprocessed images. "BWP-FBM" is the first stage of the proposed method. "Proposed" means the combination of BWP-FBM and OIS-ABM.)

Table 2
RMSE (↓) of the different algorithms in the experiments (pixels).

Experiment	Originalnal	GPTM	TPS	PLM	APAP	BWP-FBM	Proposed
Mountain (10)	3.99	0.89	1.29	0.75	0.83	**0.73**	**0.66**
Urban area (10)	139.78	4.83	0.98	**0.93**	2.66	1.02	**0.91**
Farmland (10)	1.81	1.64	1.66	1.58	1.49	**1.09**	**0.78**
Hybrid (10)	11.98	0.92	0.90	0.72	**0.48**	**0.48**	**0.42**

Note: "BWP-FBM" is the first stage of the proposed method. "Proposed" means the combination of BWP-FBM and OIS-ABM.

influenced by radiation difference and land-cover changes, as the experimental images were taken at different times, and even by different sensors. Therefore, to evaluate the transformation model more objectively, we put an emphasis on RMSE.

The RMSE values of the four experiments were calculated according to Eq. (18) and are listed in Table 2. The number of test points is listed in brackets in the leftmost column of Table 2. The "Original" RMSE was calculated to measure the initial deformation between the reference and sensed image. On the whole, the RMSE is decreased considerably by GPTM, although its outcome is not as ideal as that of the local models. As expected, TPS can eliminate the regional deformation, whereas it performs poorly in the area without enough feature points. Therefore, the RMSE in the "Mountain" area obtained by TPS is lower than that of GPTM. As shown, PLM is more robust than TPS, and it shows a higher registration precision in Table 2. Since PLM segments the image according to the extracted feature points, it is unaffected by the local absence of feature points. The first stage of the proposed method (BWP-FBM) is similar to APAP, whereas the accuracy is higher than that of APAP. When distributing the weight to each feature point, APAP uses a Gaussian function rather than IDW in the BWP-FBM algorithm. In addition, the reference point of the different blocks locates at the block top left in APAP, but it is the block center in BWP-FBM. Ultimately, the highest accuracy is obtained by the proposed method, by combining the respective advantages of the feature- and area-based methods, and locally estimating the transformation model.

On the whole, two conclusions can be drawn. One is that the proposed algorithm produces a precise registration result with BWP-FBM by estimating the transformation model block by block. The other is that OIS-ABM is able to obtain an optimized registration accuracy.

4. Conclusions

Image registration is a critical step when analyzing or processing two or more images. Given the inconsistent deformation caused by terrain relief and the degradation factors, a two-step algorithm combining and localizing the feature- and area-based methods is adopted to align remote sensing images in this paper. To obtain a relatively precise alignment, the block-weighted projective model is constructed by dividing the reference and sensed image into blocks, and estimating the model for each block with IDW. This result is then taken into the second stage, in which Huber estimation is integrated with ST to construct the outlier-insensitive model. In OIS-ABM, the linear polynomial transformation model is estimated at a finer scale. The optimal geometrical relationship is iteratively refined to rectify the BWP-FBM outcome, for a more precise result. For the four pairs of realistic images used in the experiments covering varied terrain features, the results qualitatively and quantitatively demonstrate that the proposed algorithm can achieve a reliable registration accuracy.

However, there are still some issues that need to be addressed in the future. Firstly, the local deformations are modeled artificially within small blocks, so the ground truth may be segmented into several blocks. An automatic blocking strategy should therefore be considered. Secondly, the proposed method is robust for simple and small radiation differences. As for complicated and large radiation differences, the nonlinear function could take effect.

Acknowledgements

The work was supported by the National Key R&D Program of China under Grant No. 2017YFA0604402, the National Natural Science

Foundation of China (NSFC) under Grant Nos. 61671334, 41701394, the Hubei Natural Science Foundation under Grant No. 2017CFB189.

References

Wong, A., Clausi, D.A., 2007. ARRSI: automatic registration of remote-sensing images. IEEE Trans. Geosci. Remote Sens. 45 (5), 1483–1493.

Shen, H., Meng, X., Zhang, L., 2016. An Integrated framework for the spatio-temporal-spectral fusion of remote sensing images. IEEE Trans. Geosci. Remote Sens. 54 (12), 7135–7148.

Duran, J., Buades, A., Coll, B., Sbert, C., Blanchet, G., 2017. A survey of pansharpening methods with a new band-decoupled variational model. ISPRS J. Photogramm. Remote Sens. 125, 78–105.

Li, X., Hui, N., Shen, H., Fu, Y., Zhang, L., 2015. A robust mosaicking procedure for high spatial resolution remote sensing images. ISPRS J. Photogramm. Remote Sens. 109, 108–125.

Jiang, Y., Xu, K., Zhao, R., Zhang, G., Cheng, K., Zhou, P., 2017. Stitching images of dual-cameras onboard satellite. ISPRS J. Photogramm. Remote Sens. 128, 274–286.

Gong, J., Ma, G., Zhou, Q., 2008. A review of multi-temporal remote sensing data change detection algorithms. Protein Expr. Purif. 82 (2), 308–316.

Zeng, C., Shen, H., Zhang, L., 2013. Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method. Remote Sens. Environ. 131, 182–194.

Zitová, B., Flusser, J., 2003. Image registration methods: a survey. Image Vis. Comput. 21 (11), 977–1000.

Xu, X., Li, X., Liu, X., Shen, H., Shi, Q., 2016. Multimodal registration of remotely sensed images based on Jeffrey's divergence. ISPRS J. Photogramm. Remote Sens. 122, 97–115.

Barnea, D.I., Silverman, H.F., 1972. A class of algorithms for fast digital image registration. IEEE Trans. Comput. C-21 (2), 179–186.

Cole-Rhodes, A.A., Johnson, K.L., LeMoigne, J., Zavorin, I., 2003. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. IEEE Trans. Image Process. 12 (12), 1495–1511.

Kern, J.P., Pattichis, M.S., 2007. Robust multispectral image registration using mutual-information models. IEEE Trans. Geosci. Remote Sens. 45 (5), 1494–1505.

Chen, H.m., Arora, M.K., Varshney, P.K., 2003. Mutual information-based image registration for remote sensing data. Int. J. Remote Sens. 24 (18), 3701–3706.

A. Colerhodes, I. Zavorin, J.L. Moigne, Mutual information as a similarity measure for remote sensing image registration, In: Proc. SPIE Aerosense 2001, Geo-Spatial Image and Data Exploitation II, vol. 4383, 2001, pp. 51–61.

Hel-Or, Y., Hel-Or, H., David, E., 2011. Fast template matching in non-linear tone-mapped images. In: International Conference on Computer Vision (ICCV), pp. 1355–1362.

Yu, L., Zhang, D., Holden, E.-J., 2008. A fast and fully automatic registration approach based on point features for multi-source remote-sensing images. Comput. Geosci. 34 (7), 838–848.

Sui, H., Xu, C., Liu, J., Hua, F., 2015. Automatic optical-to-SAR image registration by iterative line extraction and Voronoi integrated spectral point matching. IEEE Trans. Geosci. Remote Sens. 53 (11), 6058–6072.

Goncalves, H., Goncalves, J.A., Corte-Real, L., 2011. HAIRIS: a method for automatic image registration through histogram-based image segmentation. IEEE Trans. Image Process. 20 (3), 776–789.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60 (2), 91–110.

Yan, K., Sukthankar, R., PCA-SIFT: a more distinctive representation for local image descriptors. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2004, pp. 506–513.

Morel, J.-M., Yu, G., 2009. ASIFT: a new framework for fully affine invariant image comparison. SIAM J. Imag. Sci. 2 (2), 438–469.

Sedaghat, A., Mokhtarzade, M., Ebadi, H., 2011. Uniform robust scale-invariant feature matching for optical remote sensing images. IEEE Trans. Geosci. Remote Sens. 49 (11), 4516–4527.

Yi, Z., Zhiguo, C., Yang, X., 2008. Multi-spectral remote image registration based on SIFT. Electron. Lett. 44 (2), 107–108.

Bay, H., Tuytelaars, T., Gool, L.V., 2006. SURF: speeded up robust features. Comput. Vis. Image Underst. 110 (3), 404–417.

P.F. Alcantarilla, A. Bartoli, A.J. Davison, "KAZE Features". In: European Conference on Computer Vision (ECCV), 2012, pp. 214–227.

Rosten, E., Porter, R., Drummond, T., 2010. Faster and better: A machine learning approach to corner detection. IEEE Trans. Pattern Anal. Mach. Intell. 32 (1), 105–119.

Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., Fua, P., 2012. BRIEF: computing a local binary descriptor very fast. IEEE Trans. Pattern Anal. Mach. Intell. 34 (7), 1281–1298.

Hui, L., Manjunath, B.S., Mitra, S.K., 1995. A contour-based approach to multisensor image registration. IEEE Trans. Image Process. 4 (3), 320–334.

Dongseok, S., Pollard, J.K., Muller, J.P., 1997. Accurate geometric correction of ATSR images. IEEE Trans. Geosci. Remote Sens. 35 (4), 997–1006.

S.Z. Li, J. Kittler, M. Petrou, Matching and recognition of road networks from aerial images, in: European Conference on Computer Vision (ECCV), 1992, pp. 857–861.

Gioi, R.G.v., Jakubowicz, J., Morel, J.-M., Randall, G., 2012. LSD: a line segment detector. Image Process. Line 2 (4), 35–55.

T. Xiang, G. S. Xia, L. Zhang, H. Ningning, Locally warping-based image stitching by imposing line constraints, in 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 4178–4183.

Li, S., Yuan, L., Sun, J., Quan, L., 2015. Dual-Feature Warping-Based Motion Model Estimation. In: IEEE International Conference on Computer Vision (ICCV), pp. 4283–4291.

K. Lin, N. Jiang, S. Liu, L.F. Cheong, M. Do, J. Lu, Direct photometric alignment by mesh deformation, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2701–2709.

Goshtasby, A., Stockman, G.C., Page, C.V., 1986. A region-based approach to digital image registration with subpixel accuracy. IEEE Trans. Geosci. Remote Sens. 24 (3), 390–399.

Silpa-Anan, C., Hartley, R., 2008. Optimised KD-trees for fast image descriptor matching. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 1–8.

G. Liu, S. Chen, X. Zhou, X. Wang, Q. Guan, H. Yu, Combining SIFT and individual entropy correlation coefficient for image registration, In: Pattern Recognition, Berlin, Heidelberg, 2014, pp. 128–137.

Gong, M., Zhao, S., Jiao, L., Tian, D., Wang, S., 2014. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. IEEE Trans. Geosci. Remote Sens. 52 (7), 4328–4338.

Liang, J., Liu, X., Huang, K., Li, X., Wang, D., Wang, X., 2014. Automatic registration of multisensor images using an integrated Spatial and Mutual Information (SMI) metric. IEEE Trans. Geosci. Remote Sens. 52 (1), 603–615.

Goshtasby, A., 1987. Piecewise cubic mapping functions for image registration. Pattern Recogn. 20 (5), 525–533.

Zaragoza, J., Chin, T.J., Tran, Q.H., Brown, M.S., Suter, D., 2014. As-projective-as-possible image stitching with moving DLT. IEEE Trans. Pattern Anal. Mach. Intell. 36 (7), 1285–1298.

S. Liu, P. Tan, L. Yuan, J. Sun, B. Zeng, MeshFlow: minimum latency online video stabilization, in: European Conference on Computer Vision (ECCV), 2016, pp. 800–815.

Flusser, J., 1992. An adaptive method for image registration. Pattern Recogn. 25 (1), 45–54.

Goshtasby, A., 1986. Piecewise linear mapping functions for image registration. Pattern Recogn. 19 (6), 459–466.

M. Nejati, R. Amirfattahi, S. Sadri, A fast hybrid approach for approximating a thin-plate spline surface, in: 18th Iranian Conference on Electrical Engineering, 2010, pp. 204–208.

Zagorchev, L., Goshtasby, A., 2006. A comparative study of transformation functions for nonrigid image registration. IEEE Trans. Image Process. 15 (3), 529–538.

Bookstein, F.L., 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Mach. Intell. 11 (6), 567–585.

Bentoutou, Y., Taleb, N., Kpalma, K., Ronsin, J., 2005. An automatic image registration for applications in remote sensing. IEEE Trans. Geosci. Remote Sens. 43 (9), 2127–2137.

Holden, M., 2008. A review of geometric transformations for nonrigid body registration. IEEE Trans. Med. Imaging 27 (1), 111–128.

Ye, Y., Shan, J., Bruzzone, L., Shen, L., 2017. Robust registration of multimodal remote sensing images based on structural similarity. IEEE Trans. Geosci. Remote Sens. 55 (5), 2941–2958.

Han, Y., Choi, J., Byun, Y., Kim, Y., 2014. Parameter optimization for the extraction of matching points between high-resolution multisensor images in urban areas. IEEE Trans. Geosci. Remote Sens. 52 (9), 5612–5621.

Xiang, T.-Z., Xia, G.-S., Bai, X., Zhang, L., 2018. Image stitching by line-guided local warping with global similarity constraint. Pattern Recogn. 83, 481–497.

R. Feng, X. Li, W. Zou, H. Shen, Registration of multitemporal GF-1 remote sensing images with weighting perspective transformation model, in: 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 2264–2268.

Tobler, W.R., 1970. A computer movie simulating urban growth in the detroit region. Econ. Geogr. 46 (sup1), 234–240.

Lu, G.Y., Wong, D.W., 2008. An adaptive inverse-distance weighting spatial interpolation technique. Comput. Geosci. 34 (9), 1044–1055.

Hartley, R., Zisserman, A., 2003. Multiple View Geometry in Computer Vision. Cambridge University Press.

Li, X., Shen, H., Zhang, L., Zhang, H., Yuan, Q., 2014. Dead pixel completion of aqua MODIS band 6 using a robust m-estimator multiregression. IEEE Geosci. Remote Sens. Lett. 11 (4), 768–772.

Shen, H., Zhang, L., 2009. A MAP-based algorithm for destriping and inpainting of remotely sensed images. IEEE Trans. Geosci. Remote Sens. 47 (5), 1492–1502.

Afsharnia, H., Arefi, H., Sharifi, A.M., 2017. Optimal weight design approach for the geometrically-constrained matching of satellite stereo images. Remote Sens. 9 (9).

U. Köthe, Edge and junction detection with an improved structure tensor. In: 25th pattern recognition symposium of the german-association-for-pattern-recognition, 2003, pp. 25–32.

Ye, Y., Shan, J., 2014. A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences. ISPRS J. Photogramm. Remote Sens. 90, 83–95.

Y. Ye, Fast and robust registration of multimodal remote sensing images via dense orientated gradient feature, ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII-2/W7, 2017, pp. 1009–1015.

Han, Y., Bovolo, F., Bruzzone, L., 2015. An approach to fine coregistration between very high resolution multispectral images based on registration noise distribution. IEEE Trans. Geosci. Remote Sens. 53 (12), 6650–6662.

Yang, K., Pan, A., Yang, Y., Zhang, S., Ong, H.S., Tang, H., 2017. Remote sensing image registration using multiple image features. Remote Sensing 9 (6), 581 (1-21).