



ELSEVIER

Contents lists available at ScienceDirect

## Infrared Physics &amp; Technology

journal homepage: [www.elsevier.com/locate/infrared](http://www.elsevier.com/locate/infrared)

## A binary method of multisensor image registration based on angle traversal

Huaidan Liang<sup>a,b</sup>, Chenglong Liu<sup>a,b,c</sup>, Bin He<sup>a,\*</sup>, Ting Nie<sup>a</sup>, Guoling Bi<sup>a</sup>, Chang Su<sup>a,b</sup><sup>a</sup> Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China<sup>b</sup> University of Chinese Academy of Sciences, Beijing 100049, China<sup>c</sup> Key Laboratory of Airborne Optical Imaging and Measurement, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

## ARTICLE INFO

## Keywords:

Registration  
Multisensor remote-sensing images  
Descriptor  
Local grayscale inversion  
Binary

## ABSTRACT

Registration is a key step in image preprocessing. Absorbing the binary idea of BRIEF, we took all the feature points as a whole and generated a new binary descriptor that was irrelevant to the gray of the neighborhood of a feature point in the registration of visible light with infrared images. For each selected feature point, we constructed a circular area centered on it and divided the circumference equally as the sampling interval of the feature points. We then traversed all the intervals and built a binary code string as a descriptor for the feature point according to the relative position of the feature point. During the process of angle traversal, the Hamming distance was applied to identify the most similar descriptor pair. We then adopted an improved clustering algorithm to eliminate mismatched pairs. Experiments have demonstrated that this kind of descriptor has high accuracy and robustness for multisensor remote-sensing images. In particular, when dealing with the combination of visible and far-infrared images with local grayscale inversion, registration can also be accomplished well.

## 1. Introduction

Registration is the process of aligning two or more images of the same scene taken in different ways [1]. This ability is a necessary prerequisite for many practical problems, such as image fusion [2–4] and change detection [5,6]. Therefore, determining how to improve the accuracy and efficiency of registration has attracted substantial attention [7–9].

According to the different judgment basis in the registration process, existing registration methods can be roughly divided into feature-based methods, such as key points [10–12], edges or contours [13–16] and regions [17]; transform-field-based methods, such as frequency features; and statistical-based methods, such as mutual information [4,18–21]. Among these methods, feature-point-based methods are the best in terms of their stability and robustness for remote-sensing images with various scenes and high requirements. The scale-invariant feature transform (SIFT) algorithm proposed by Lowe was a classic method of feature point extraction and matching, with good scale characteristics [11]. Many improved methods based on the principle of SIFT were proposed to improve its speed and rotation performance [22]. The speeded up robust features (SURF) algorithm proposed by Bay et al. was an improved and accelerated version of SIFT that increased the speed of detecting the feature points; however, the stability of the feature points

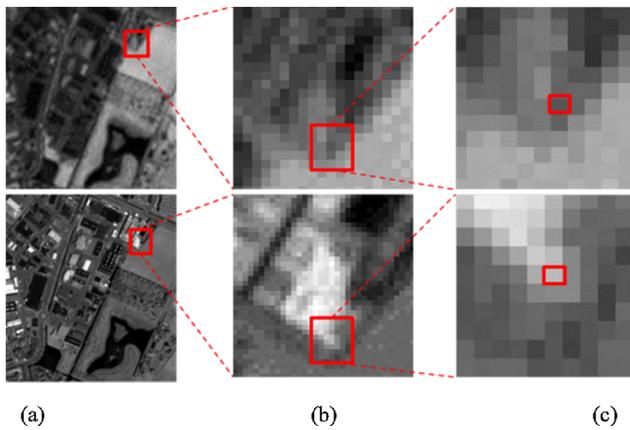
was slightly decreased [12,23].

As alternatives to methods such as SIFT and SURF, methods were created based on various principles. The binary robust independent elementary features (BRIEF) descriptor used a binary code string as a descriptor vector and the Hamming distance for matching, which greatly increased the speed of registration [24]. Because the BRIEF descriptor did not have directionality or another means to address rotation, rotation had a substantial impact on the results. Ethan Rublee et al. combined features from accelerated segment test (FAST) and BRIEF, proposing an oriented FAST and the rotated BRIEF (ORB) algorithm and adding an orientation to the FAST feature points [25,26]. The authors made the feature points rotation invariant and proposed a pyramid method to achieve scaling invariance. The binary robust invariant scalable keypoints (BRISK) algorithm was also a binary descriptor with good scale invariance and robustness, but it took more time than ORB [27].

Registration of multisensor remote-sensing images is more practical. As technology advances, different types of sensors can be applied to enhance the complementarity of information [28,29]. For example, in the field of remote sensing, visible-light sensors and infrared sensors are used in conjunction to satisfy the needs of all-day reconnaissance. With the infrared sensors detecting radiation information and the visible-light sensors detecting reflection information, combining the two types

\* Corresponding author.

E-mail address: [lianghuaidan14@mails.ucas.ac.cn](mailto:lianghuaidan14@mails.ucas.ac.cn) (B. He).



**Fig. 1.** The local grayscale inversion. Difference in gray level of the same feature point in a visible light image and an infrared image. (a) Images of the same object on the visible and infrared images. (b) Partial enlargements of the object. (c) The extracted feature points.

of images can enhance the complementarity of the scene.

The difficulty of multisensor image registration lies in the fact that the image outputs by different sensors have different grayscale characteristics, even with local gray inversion between the different kinds of images. This is because the visible light sensor detects a band of 400–700 nm, which is related to the reflection of the object's external illumination; while the long-wave infrared sensor detects a band of 8000–14000 nm, which is related to the temperature and material of the object. The grayscale of pixels around the feature points varies greatly. There may be large errors in the gradient direction of the feature point according to pixels in its neighborhood or sampling in its neighborhood when the classical methods are used to describe or match the feature points.

As shown in Fig. 1, a white object has a high level of brightness in a visible-light image. However, in an infrared image, its radiation intensity may be weak due to its low absorption of sunlight and its low surface temperature. As a result, the gray level is much lower in the far-infrared image. The gray gradient directions of the same object in the two images are opposite. In this case, according to the sampling method of the classical method, mis-matching occurs. An important direction for future research is to study methods that are independent of the characteristics of remote-sensing images and are suitable for the registration of different data sources.

At the same time, many practical applications have high requirements for speed in remote-sensing image registration. Many methods consume much time in the process of feature extraction and matching

and cannot meet the requirements of real-time registration tasks. With the increasing spatial resolution and spectral resolution of remote-sensing images, the data of remote-sensing images will increase, and the data processing time will be longer. In this case, how to ensure the accuracy of registration and how to improve the speed of registration are equally important. The binary idea of the BRIEF algorithm can accelerate the registration observably. However, the BRIEF algorithm does not work well in multisensor remote-sensing images. Thus, we designed a new matching method based on its binary principle according to the characteristics of remote-sensing images.

On the one hand, this method discards the traditional method, which is based on the gray levels of the neighborhood of feature points and can be made to work with the registration of multisensor remote sensing images; on the other hand, it ensures the correct rate through optimal matching and reduces errors of pixel coordinates resulting from scale differences by global matching. In this way, it is possible to effectively avoid the interference of the mismatched points in the subsequent processing steps, thereby avoiding manual intervention and greatly saving on time.

## 2. Materials and methods

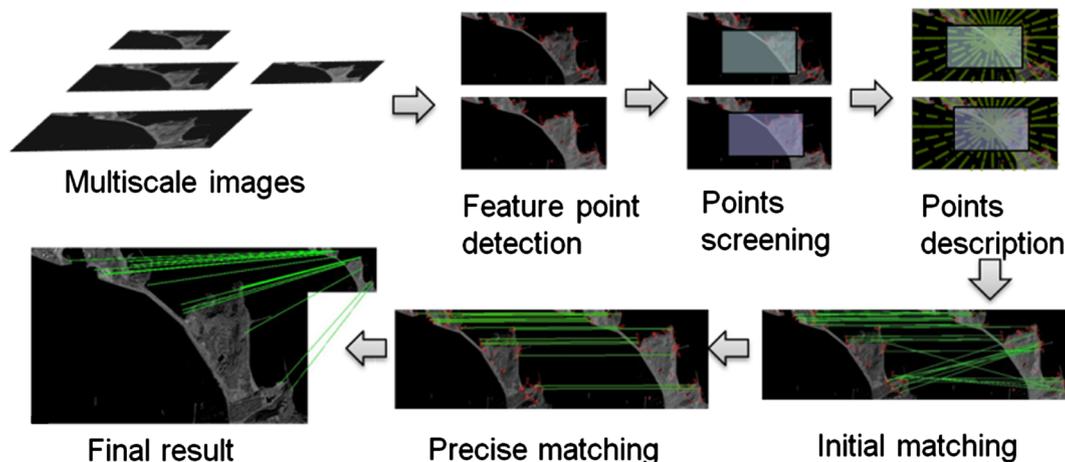
A flow chart of the proposed method is illustrated in Fig. 2. Each step will be explained in detail.

### 2.1. Feature point detection

Most of the multisensor remote-sensing images differ greatly in spatial resolution. For the same image, the feature points extracted will change with the resolution changing, as shown in Fig. 3. Only a small number of feature points is necessary for the registration of low-resolution images. The feature points can be extracted from the low-resolution image after the image pyramid is built to decrease the computing time.

Based on the existing methods of feature point extraction, the methods of description and matching are studied in this paper. We select SURF as our algorithm to extract feature points in consideration of both accuracy and speed. We set a threshold for SURF to select a certain number of the strongest feature points according to the value of the Hessian matrix, which is the local maximum determinant for a given point, as defined in Eq. (1). This paper sets a higher threshold when extracting feature points from the image; only the particularly stable feature points remain.

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (1)$$



**Fig. 2.** The flow chart of the proposed method.

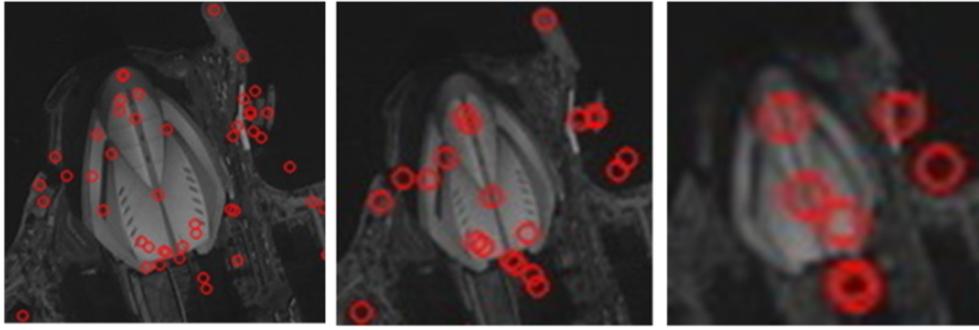


Fig. 3. The points extracted by SURF with different resolution.

where  $L_{xx}(x, y, \sigma)$  is the convolution of the Gaussian second-order derivative  $\frac{\partial^2 g(y)}{\partial x^2}$  with image  $I$  at point  $(x, y)$  and scale  $\sigma$ .  $L_{xy}(x, y, \sigma)$  and  $L_{yy}(x, y, \sigma)$  are similar to  $L_{xx}(x, y, \sigma)$ . SURF adopts the convolution of the boxfilter with the integral image  $D_{xx}$ ,  $D_{xy}$  and  $D_{yy}$  to approximate  $L_{xx}$ ,  $L_{xy}$  and  $L_{yy}$ , respectively, to speed up the calculation. The determinant of the Hessian matrix is computed by Eq. (2).

$$\det(H) = D_{xx}D_{yy} - (0.9D_{xy})^2 \tag{2}$$

where 0.9 is set in accordance with practice.

Due to the large variation in spatial resolution of a single pixel in each group, we implement the method of building image pyramids to conduct the experiments. This paper adopts the second bilinear interpolation in downsampling. By reducing the definition of the images to be matched, extracting the representative feature point pairs, and calculating the positions of these points in the original image, we can match the points in the low-definition image with the points in the high-definition image using an improved image scale upgrade, as shown in Fig. 4. We improved the upgrade to limit the uncertainty as low as possible in the upgrade process by constructing a pair of similar triangles, which are formed by the three furthest points in the low-definition image.

Then, we obtain the locations of the feature points.

### 2.2. Points screening

We set a window for each image to screen the feature points according to their positions to obtain the feature points in the central area of the images. We define the feature points in the window as the main feature points. The screening process is shown in Fig. 5. Every point in Fig. 5 is a feature point, and the big points are the points in the window, defined as the main feature points. The subsequent matching process is performed between the main feature points. Windowing can reduce the amount of computation as well as the impact of distortion on the edges of images.

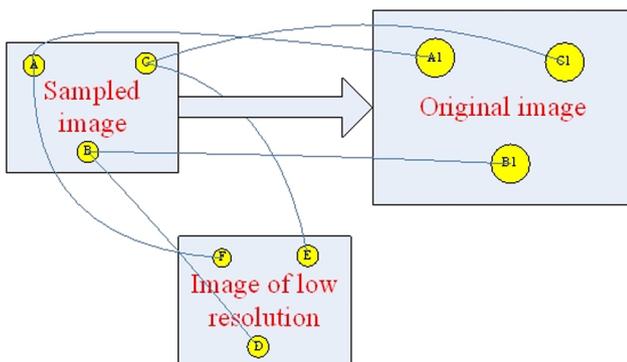


Fig. 4. The strategy from rough to fine.

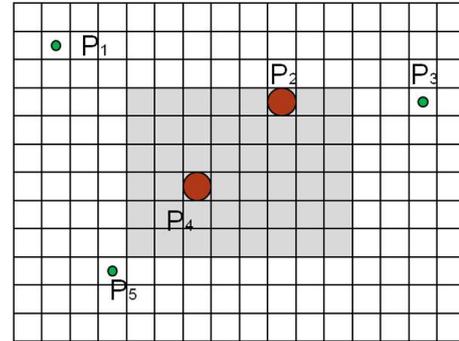


Fig. 5. Set a window to screen the points.

### 2.3. Generation of descriptor

First, we construct vectors from one main feature point to all the remaining feature points one by one. A polar coordinate system is constructed for each main feature point, and its circumference is divided into several equal angular intervals. The angle interval,  $\gamma$ , can be called a scanning step during angle traversal, as shown in Eq. (3), where  $n$  is the total number of angular intervals. The angle interval is recommended to be set smaller than  $3^\circ$  to ensure accuracy. Moreover, the total number,  $n$ , should be increased to enhance the distinction of the descriptors.

$$\gamma = \frac{360^\circ}{n} \tag{3}$$

Second, we scan from  $0^\circ$  of each main feature point in the anticlockwise direction to check if there are other feature points within the angle intervals. Downward rounding is executed to ensure the absence of  $360^\circ$ . That is, for example, when the sampling intervals are set at  $1^\circ$ , the angles can be limited to  $0\text{--}359^\circ$ , 360 intervals in total. The anticlockwise direction is the direction of increasing angle.

Eq. (4) is set to calculate the angles related from one main feature point to the remaining feature points. The operator  $[ ]$  represents downward rounding. Then, we can obtain a set of angles  $\{\beta\}$ .

$$\beta_i = \begin{cases} \left[ \arctan \frac{P_{main}(y) - P_i(y)}{P_{main}(x) - P_i(x)} \right], & P_{main}(y) > P_i(y) \\ \left[ \arctan \frac{P_{main}(y) - P_i(y)}{P_{main}(x) - P_i(x)} + 360 \right], & P_{main}(y) < P_i(y) \end{cases} \tag{4}$$

where  $P_{main}(x)$  is the horizontal coordinate of the main point, and  $P_i(x)$  is the horizontal coordinate of the remaining feature points.

Then, the value of the corresponding bit in the descriptor is transformed according to Eq. (5), when the angle interval,  $\gamma$ , is set not equal to 1.

$$\alpha_i = [\beta/\gamma] \tag{5}$$

In this way, when there are other feature points in one angle interval, the corresponding bit is marked as 1; when there is no feature

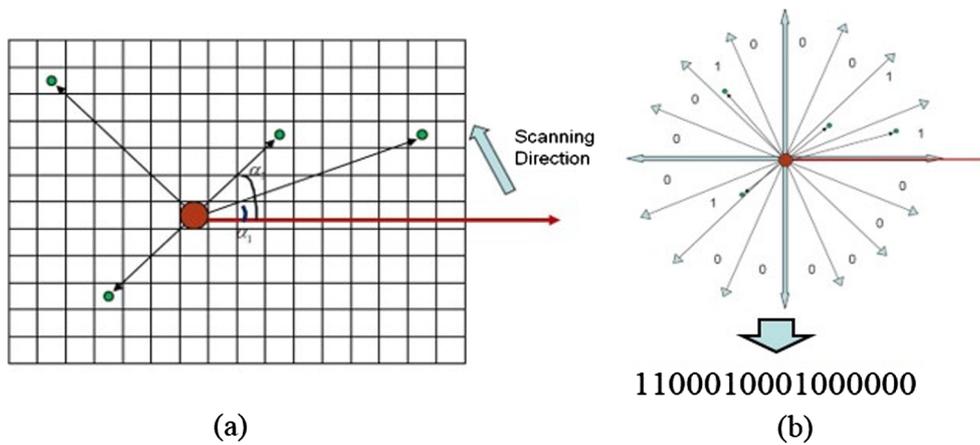


Fig. 6. (a) Constructing vectors and scanning; (b) an example of the generation of a descriptor.

point, it is recorded as 0, as shown in Eq. (6).

$$\alpha_i = \begin{cases} 0, & \text{there is no point} \\ 1, & \text{otherwise} \end{cases} \quad (6)$$

That is, a bit  $\alpha$  indicates whether feature points exist in the current angle interval. We obtain a set of binary code strings composed of  $\alpha$  after traversing all the angle intervals. This is the method used to complete angle traversal. This code string is used to describe one main feature point and is defined as the descriptor of this point. One image is set as the reference image, and the image to be registered is set as the floating image. A descriptor is generated for each main feature point in the two images.

The length of the descriptor can be dynamically adjusted according to the definition of images and accuracy requirements by adjusting  $n$ . The generation process of a descriptor and an example are illustrated in Fig. 6. For clarity, the angle interval in Fig. 6 is large. According to the positions of the feature points in Fig. 6, the descriptor of this point is **1100010001000000**. The descriptors of all the main feature points are generated in both images in this way. The total quantity of all the main feature points is not large, as a result of windowing.

#### 2.4. Matching process of the descriptor

The matching process is conducted based on the principle of minimum distance. The process is shown in Fig. 7. First, we calculate the Hamming distance between two descriptors, where one comes from the reference image and the other comes from the floating image. Then, we shift all the descriptors from the floating image by one bit while all the descriptors from the reference image are invariant and recalculate the Hamming distance. This operation is repeated to compare all the combinations of main feature points in the window and seek the minimum distance until the shifting threshold is reached, as illustrated in Fig. 7. A descriptor is taken from the reference image and the registration image, respectively, to calculate the Hamming distance between them.

In this paper, we adopt the operation “XOR” and count the number of 1s. The sum is the Hamming distance. The formula is given in Eq. (7).

$$D_H(x, y) = \sum_{i=0}^{k-1} \sum_{j=0, j \neq i}^{k-1} \alpha_{ij} \quad (7)$$

In the case of this direct comparison, the original descriptors apply only in the absence of rotation. The code string of a descriptor needs to be shifted to calculate the rotated angle. When the image is rotated, the descriptors of the rotated image still start at  $0^\circ$  in the horizontal direction of the new image. At this moment, all of the feature points are rotated by a certain angle as a whole. Accordingly, our descriptor

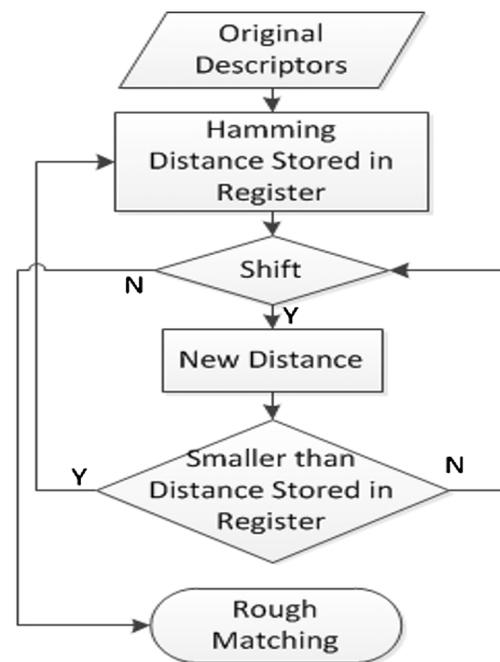


Fig. 7. The process of matching.

becomes 0011000100010000, as shown in Fig. 8.

The descriptor can digitally simulate the image rotation process by shifting its bits head-to-tail to solve the image rotation problem. This descendent process of the descriptor shifting from end to beginning is shown in Fig. 9. We can conclude that the rotated angle is equal to two angle intervals in Fig. 9.

When we divide the circumference into intervals of  $1^\circ$ , the length of the descriptor bits is 360. Moving the furthest-right bit to the beginning to obtain a new code string corresponds to rotating the image anticlockwise by  $1^\circ$ . Another Hamming distance needs to be calculated after a single shift. We can then repeatedly calculate the Hamming distance to find the pair with the smallest distance.

If the Hamming distance after the shift is greater than that before the shift, the matching error increases and the extent of matching is weakened. By contrast, the match is not necessarily the best if the Hamming distance after the shift is less than that before the shift. In the case that the shift runs into a certain angle, the Hamming distance decreases suddenly to a small value. The process is similar to a disc composed of sectors that rotate exactly about the center so that only one position is perfectly matched. As shown in Fig. 10, if each color represents 0 or 1, the best match for both disks is achieved only when

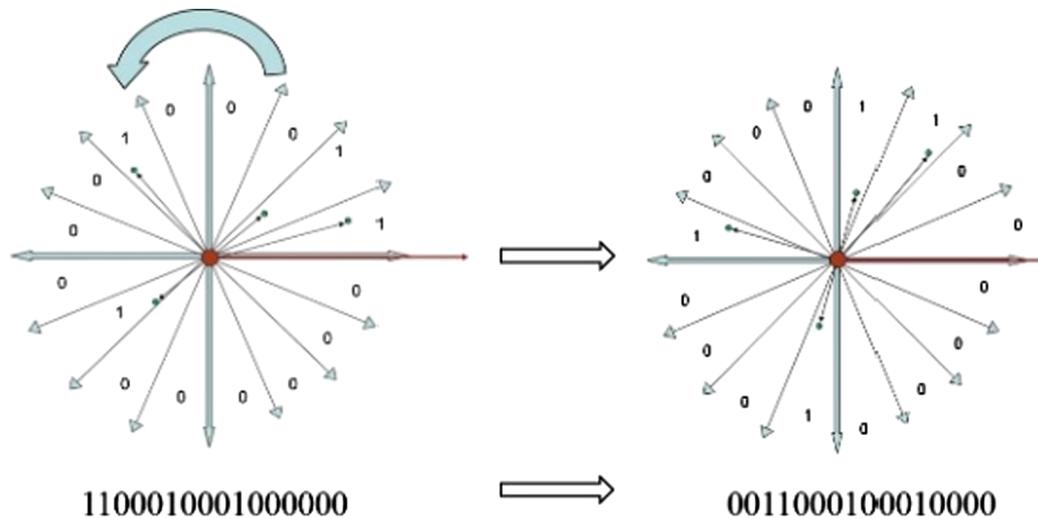


Fig. 8. The process of rotation.

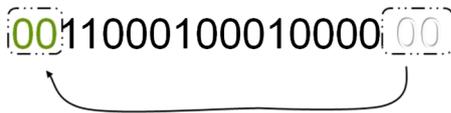


Fig. 9. The descriptor shifting from end to beginning.

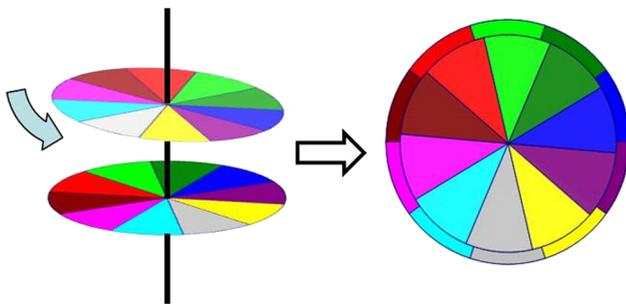


Fig. 10. The best match for two disks is achieved at a specific angle.

one is rotated to a certain angle while the other is static. For the two images, the extent of matching is the best when the Hamming distance between the pair of main feature points is the minimum. The number of bits that shifted indicates how many angular intervals the image has rotated, i.e., the angle of rotation between the two images.

After a circumferential or specific rotation, we can obtain the minimum Hamming distance along with the corresponding point-pair, the descriptors that achieve the minimum, and the amount of the shift. We take this pair of feature points from the two images as the best homologue points, representing the first and best choice for the matching process. Their corresponding positions in coordinates can be set as the original points and can be used to calculate the translation parameters. The product of the amount of the shift and the angle interval represents the rotated angle of the two images. Since the angle intervals can be set according to one’s purpose, the descriptor in this paper can achieve high accuracy and recognition of the rotated angles and keep the matching performance stable at any rotated angle. This is superior to the method of the dominant orientation. In this way, the descriptor is irrelevant to the gray levels or intensity in the neighborhood of the feature point.

Then, the two descriptors with the highest extent of matching adopt the bitwise “AND” operation. In this way, we can obtain the intersection of the feature points in the two images. The 1s remaining after the AND operation represent feature points that exist in both the corresponding angle intervals in the two images. Based on these 1 bits, these

feature points are regarded as initial homologue points. The remaining feature points are eliminated because they have no counterpart with which to form a pair. We can obtain a roughly matched table. In this table, we list the remaining point pairs in a single column one by one.

This paper sets a higher threshold when extracting feature points, and the feature points are screened by the window so that the number of feature points is effectively controlled. Even if we compare all combinations, it will not take too much time.

### 2.5. Eliminate the mismatched pairs

As the feature points in the same angle interval are taken as homologue points, there may be one point corresponding to a few points in the roughly matched table. Therefore, we first need to address the multicombinations to ensure that the mapping relation is unique.

First, the column is split by copying one point to form several new pairs to ensure that the mapping relation is unique in the roughly matched table. A small number of mismatched pairs may exist, and we apply statistical methods to solve this problem. Considering that the current matched table is a set of pairs of points in two-dimensional space, we refer to the basic idea of the Fisher discriminant to find the best and most easily classified projection method to address multi-dimensional data [30]. Therefore, we project the matched table in a straight line to form a one-dimensional space or set. The Fisher criterion function is defined in Eq. (8).

$$J_F(w) = \frac{(m_1 - m_2)^2}{S_1^2 + S_2^2} \tag{8}$$

where  $(m_1 - m_2)$  is the difference between the means of the two vectors;  $S_1^2$  is the in-class dispersion, which can be selected as the variance; and  $w$  is the projection direction.

To begin, we implement a random sample consensus (RANSAC) method to seek the best homologue points. However, experiments have shown that RANSAC is not satisfactory when the definition of an image is not high or the feature points are densely distributed because this method requires a higher number of correctly matched pairs. Therefore, the table must be processed to increase the degree of differentiation.

Clearly, to increase the degree of differentiation, i.e., to find the  $w$  that makes  $J_F(w)$  maximized as the projection direction, the numerator of  $J_F(w)$  must be as large as possible, and the denominator must be as small as possible. However, the formula above does not contain  $w$ , so  $J_F(w)$  can be rewritten as Eq. (9):

$$J_F(w) = \frac{W^T S_b W}{W^T S_w W} \quad (9)$$

where  $S_b$  is the dispersion matrix between samples, and  $S_w$  is the dispersion matrix within the total sample. The Lagrange multiplier method can be used to calculate the  $w$  to maximize  $J_F(w)$ . After calculating the partial derivatives, we obtain the final  $W^*$ , as shown in Eq. (10). Here, we do not show the process in unnecessary detail. We are concerned with the form of Eq. (10). If we want to increase the degree of differentiation,  $(m_1 - m_2)$  should be multiplied.

$$w^* = S_w^{-1}(m_1 - m_2) \quad (10)$$

With reference to this transformation, we first determine the distances between all the feature points and the main point in the matched table and then calculate the ratios of the two distances as  $R_1$ ,  $R_2$  and so on. Then, the dispersion of the ratios is determined in the form of the standard error. In this way, we transform the Euclidean distance for each pair of feature points in the pairs into the set  $\{R_1, R_2 \dots\}$ . The final ratio  $R$  represents the scaling parameter  $S$  between the two images. We set the threshold for dispersion based on Eq. (10). Similar to the normalization method, a scale factor  $S_w$  is used to regulate the degree of differentiation. As shown in Eq. (11), we can transform the set  $\{R_1, R_2 \dots\}$  into a normalized one  $\{r_1, r_2 \dots\}$  by the factor  $S_w$ .

$$r_x = S_w \times R_x \quad (11)$$

We use the least squares (LS) method to address the set roughly and to obtain the approximate results. Then, we use K-means clustering to eliminate mismatched point pairs and obtain precise results.

Although some feature-point extraction methods can achieve sub-pixel accuracy, we have included some margin in the filtering process in consideration of the systematic errors caused by the discreteness of digital images. Deviation from the reference  $r^*$  by 0.5%, which allows for a one-pixel deviation in an image of size  $1K \times 1K$ , is permissible. Therefore, the distance of the clustering algorithm is limited to 1% as the discriminant clustering threshold  $d$ . The reference  $r^*$  is calculated via the K-means clustering algorithm in a following paragraph. The LS filter residual error  $L$  is a relatively large interval containing  $R$ . The threshold  $d$  is set as the length of the interval to calculate  $k$ , as shown in Eq. (12). The parameter  $k$  is the initial number of centers of clusters. We limit  $d$  smaller than 1% to ensure the correct rate, even if some correct pairs may be eliminated due to this strict criterion.

$$k = L/d \quad (12)$$

At this time, we can obtain all the parameters applied in the transform.

### 3. Experiments

#### 3.1. Conditions

Experiments were conducted to verify the performance of the proposed method and to compare it with other methods. The methods used for comparison were the standard SURF and ORB. The same machine was used for the test in all cases (proposed, SURF and ORB) with Visual Studio 2012 as the experimental environment. In the experiments, the

**Table 1**  
Information about the experimental images.

Group	Definition	Spatial resolution	Image type
First	3436 * 3232	0.8 m	visible-light
	289 * 288	10 m	far-infrared
Second	3340 * 3340	0.8 m	visible-light
	228 * 284	10 m	far-infrared
Third	3754 * 3324	0.8 m	visible-light
	330 * 312	10 m	far-infrared

angle interval was set to  $1^\circ$ , and the size of the screening windows was 60% of that of the original images. During the experiments conducted in this paper, we downsampled the visible-light images to obtain pyramid images with different spatial resolutions.

Six groups of typical visible-light and far-infrared remote-sensing images of different scenes were chosen in this experiment. Three groups are illustrated in detail to compare their validity.

The first group of experimental images is the airport area. The second group of experimental images is the lake area. The third group of images is the harbor area. In the third group of images, there is a large angle of rotation between the two images. Information about the three groups of experimental images is shown in Table 1. The feature points extracted by ORB and SURF are not limited as shown in Fig. 11 and Fig. 12. The threshold of the ORB was the recommended value of 64. In the SURF algorithm, the threshold of the Hessian matrix was the default value of 300. The points adopted in the proposed method were extracted by SURF with higher threshold values. Furthermore, the thresholds of the Hessian matrix used in proposed method were different. In the visible images, the matrix threshold was 7000, and in the far-infrared image, the matrix threshold was 2000.

#### 3.2. Results and discussion

The results of the first two groups are also illustrated in Figs. 11, 12 and 13. The matching processes of ORB and SURF were based on a point-to-point method. The possibility of accomplishing the registration increased as more points were extracted. However, the proposed method did not require many points. Using fewer points than ORB and SURF, the proposed method could achieve more pairs.

The numbers of points extracted and matched are listed in Table 2. In the first and third groups, the matched pairs accomplished by ORB and SURF were all incorrect. Even when we set a stricter threshold for the distance for ORB and SURF, the matched pairs with the least distance were still incorrect. In the second group, all matched pairs accomplished by ORB were still incorrect, while pairs accomplished by SURF were partially correct. It could be seen that the proposed method had a stable correct rate, which was very important in automation and practical application.

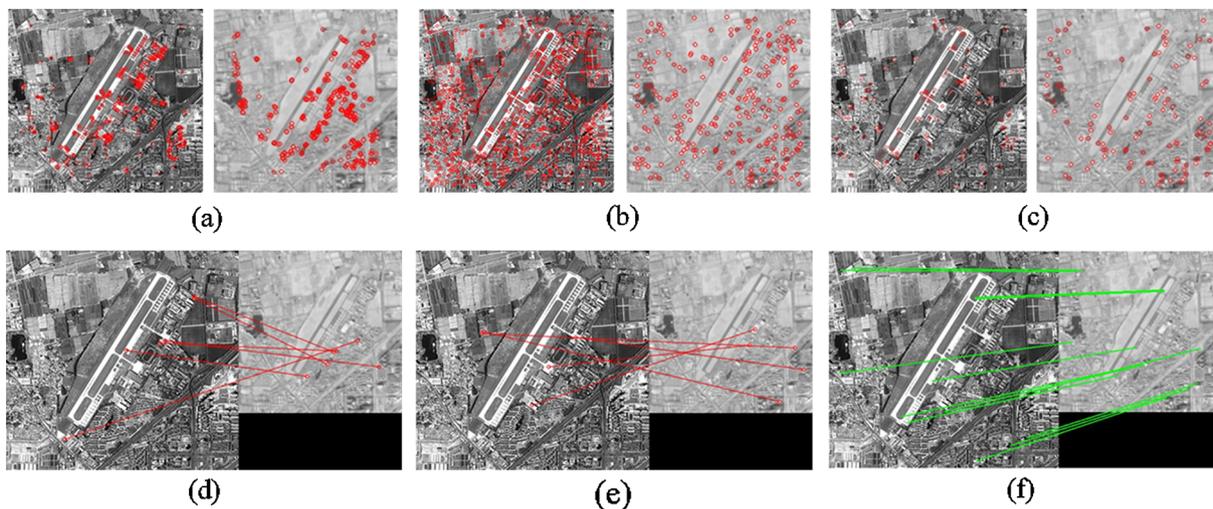
Regarding the matching speed, we could see that the time consumed by the proposed method was approximately equal to that spent by ORB, even in the case of large angle rotation.

The overlapping images obtained after the registration process are shown in Fig. 14. The rectangle boxes in the overlay represent the portions with good registration. The red<sup>1</sup> squares show their enlarged details.

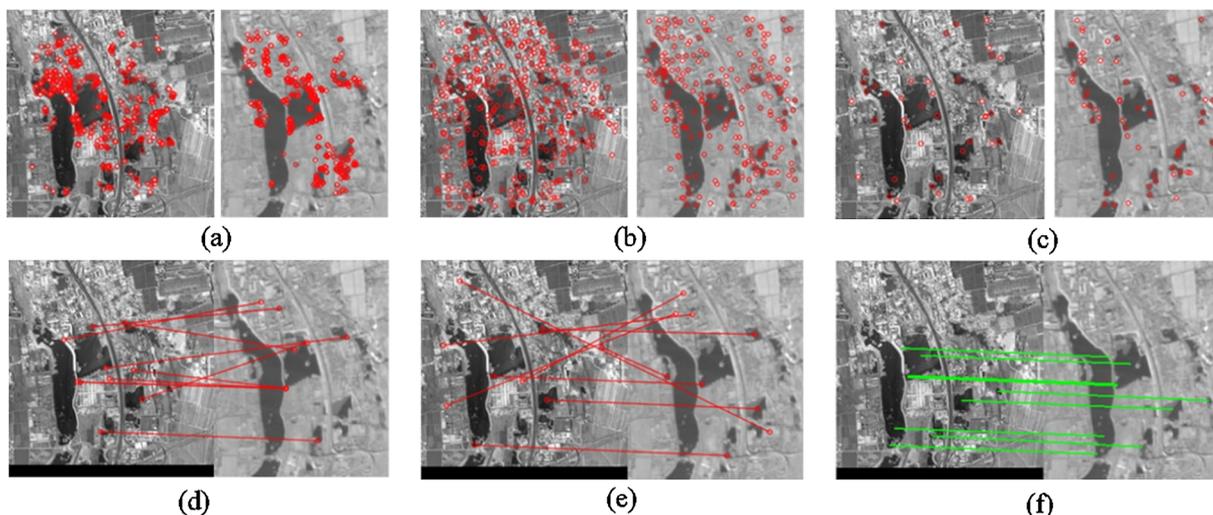
We also present a comparison chart of the matching results for three other regions in Figs. 15, 16 and 17. It can be seen that the proposed algorithm performs significantly better than ORB and SURF in the registration of visible-light and far-infrared images. The numbers of extracted and matched points are listed in Table 3.

Fig. 18 shows the overlapping images obtained after the registration process. The six pairs of images displayed in the experiment are visible-light and far-infrared images, respectively. It can be clearly seen that the same scene between two images of each group has obvious differences in gray levels even with local gray inversion. These obvious gray-level differences lead to the failure of the strategy of description and matching based on the dominant direction of feature points. The dominant orientation is the direction of the maximum value corresponding to the histogram of the directional distribution formed by counting the gradation gradients of a fixed number of pixels in the neighborhood of the feature point. The details of the dominant direction can be found in SIFT-like algorithms [15]. As a result, SURF and

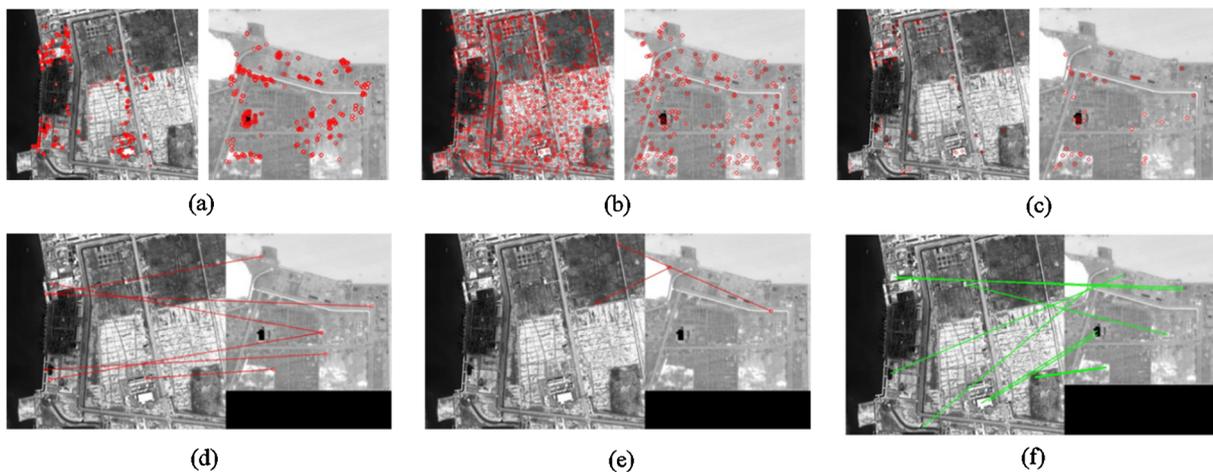
<sup>1</sup> For interpretation of color in Fig. 14, the reader is referred to the web version of this article.



**Fig. 11.** The result of the first group. (a) Feature points extracted by ORB in visible-light and far-infrared images. (b) Feature points extracted by SURF. (c) Feature points adopted in the proposed method that are extracted by SURF with a higher determinant of the Hessian matrix. (d) The mapping between the sampled visible-light image with the far-infrared image accomplished by ORB. (e) The mapping accomplished by SURF. (f) The mapping accomplished by the proposed method.



**Fig. 12.** The result of the second group. (a) Feature points extracted by ORB in visible-light and far-infrared images. (b) Feature points extracted by SURF. (c) Feature points adopted in the proposed method that are extracted by SURF with a higher determinant of the Hessian Matrix. (d) The mapping between the visible-light image with the far-infrared image accomplished by ORB. (e) The mapping accomplished by SURF. (f) The mapping accomplished by proposed method.



**Fig. 13.** The result of the third group. (a) Feature points extracted by ORB in visible-light and far-infrared images. (b) Feature points extracted by SURF. (c) Feature points adopted in the proposed method that are extracted by SURF with a higher determinant of the Hessian Matrix. (d) The mapping between the visible-light image with the far-infrared image accomplished by ORB. (e) The mapping accomplished by SURF. (f) The mapping accomplished by the proposed method.

**Table 2**  
Results of the experimental images.

Group	Method	Points extracted in visible-light image	Points extracted in infrared image	Matched pairs	Correct pairs	Time consumed (ms)
First	ORB	500	426	6	0	35.04
	SURF	1075	251	5	0	333.67
	Proposed	126	110	11	11	33.29
Second	ORB	467	412	9	0	32.96
	SURF	504	230	9	3	328.81
	Proposed	62	43	8	8	32.14
Third	ORB	382	219	6	0	33.18
	SURF	1151	199	2	0	330.24
	Proposed	69	37	8	8	33.78

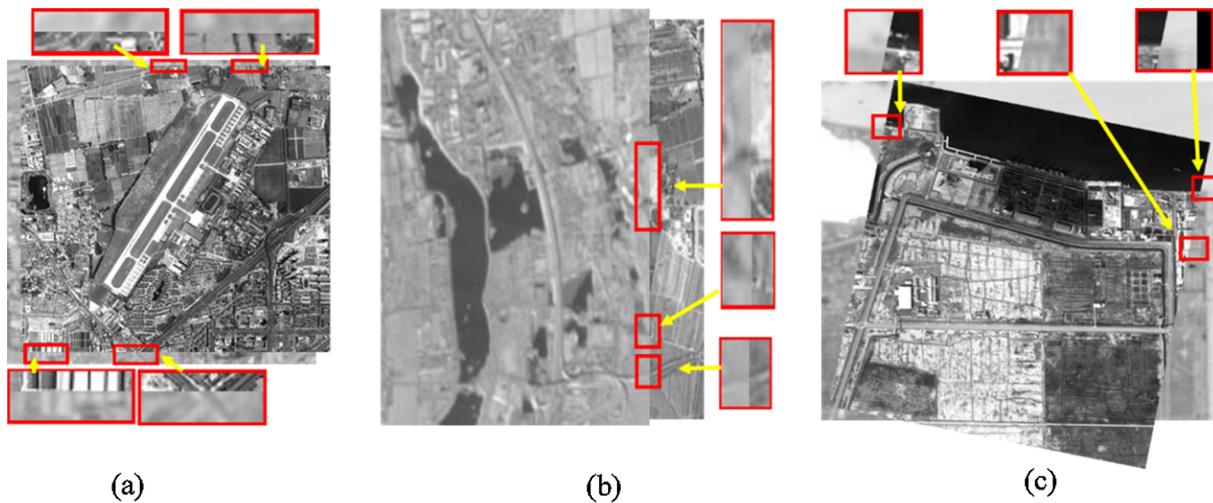
ORB algorithms are likely to fail or mismatch in the experiment. However, the proposed method can still complete the registration. This method uses the idea of a global description of feature points. Judging from the relative positions of the feature points helps eliminate

the effects of local gray changes. Even if the gray-level inversion between the far-infrared image and the visible-light image occurs, as long as the homologous points have strong intensity and can be extracted simultaneously, the calculation process will not be affected. Therefore, the proposed method is insensitive to phenomena such as gray inversion or light variation.

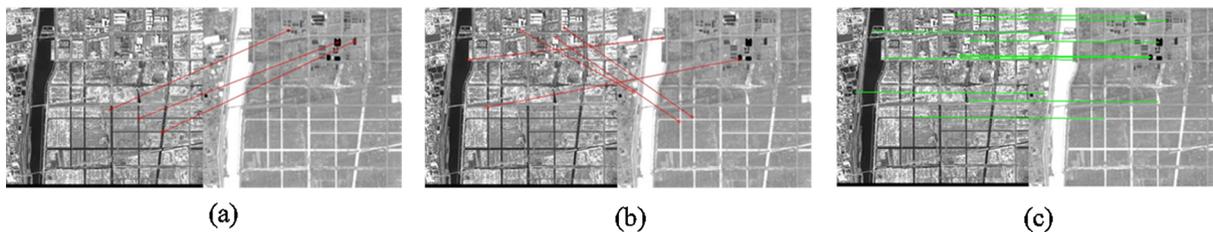
Methods based on typical features, such as edges or contours [13–16] and regions [17], are robust to geometric variations, occlusion, background clutter and noise. However, they are highly sensitive to structural differences caused by unimportant structures. This sensitivity results in a severe drop in matching performance when large contrast differences occur in the input image. Compared with the typical feature-based method, this paper can extract more powerful common features from different modalities and is less sensitive to contrast differences.

The method proposed in this paper has two significant advantages. First, the feature points are described by a string of bits; thus, the calculation is quick. Second, robust results can be achieved for images that differ substantially with respect to gray levels.

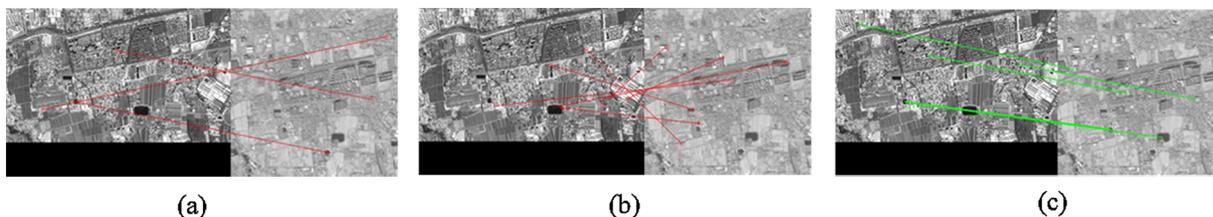
The reason we chose the SURF algorithm to extract the feature point rather than FAST is that we find that the point extracted by FAST is



**Fig. 14.** The overlapping images of the original visible-light images and the far-infrared images. (a) The overlapping image in the first group. (b) The overlapping image in the second group. (c) The overlapping image in the third group.



**Fig. 15.** The mapping of the fourth group between the sampled visible-light images with the far-infrared images. (a) The mapping accomplished by ORB. (b) The mapping accomplished by SURF. (c) The mapping accomplished by the proposed method.



**Fig. 16.** The mapping of the fifth group between the sampled visible-light images with the far-infrared images. (a) The mapping accomplished by ORB. (b) The mapping accomplished by SURF. (c) The mapping accomplished by the proposed method.

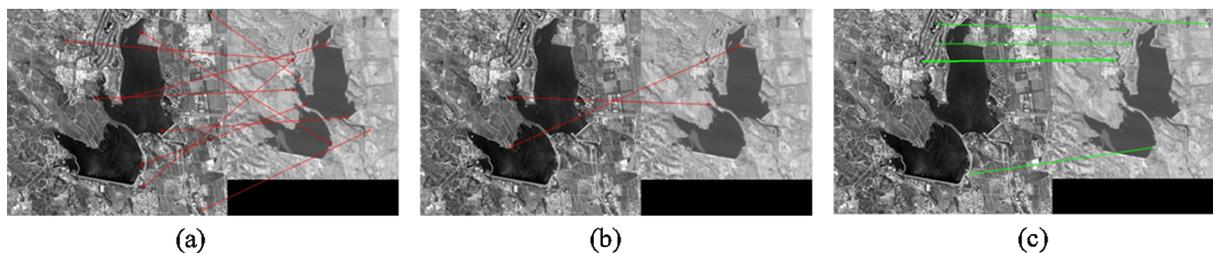


Fig. 17. The mapping of the sixth group between the sampled visible-light images with the far-infrared images. (a) The mapping accomplished by ORB. (b) The mapping accomplished by SURF. (c) The mapping accomplished by the proposed method.

Table 3  
Results of the experimental images.

Group	Method	Points extracted in the visible-light image	Points extracted in the infrared image	Matched pairs	Correct pairs	Time consumed (ms)
Fourth	ORB	413	395	3	0	31.76
	SURF	985	301	5	0	346.18
	Proposed	126	110	9	9	29.83
Fifth	ORB	436	378	3	0	33.21
	SURF	940	221	7	0	338.59
	Proposed	62	43	4	4	31.36
Sixth	ORB	458	336	9	2	34.97
	SURF	1327	232	2	1	345.85
	Proposed	97	58	5	5	33.13

more likely to deviate from the location it should be. This issue may be attributed to the criterion of FAST being simplified. In contrast, SURF adopts the integral image to extract points. The location calculated by SURF is consecutive and can achieve the subpixel accuracy.

During practical registration, the shift operation does not have to traverse 360°. For example, the angles between the visible-light load and infrared load may be only a few degrees in the same system. Therefore, the shift operation can be completed by moving a few bits to the left or to the right, which is equivalent to rotating one type of image left or right to make the two image types coincident.

Moreover, the lengths of the descriptor can be increased by decreasing the interval to improve the angle resolution, as shown in Eq.

(3). In The parameters can be dynamically adjusted to satisfy different needs in the practical process based on the requirements.

### 4. Conclusions

This paper presents a method to describe feature points based on the distributions of feature points to solve the problem of registration in the field of far-infrared and visible-light images. Then, the method to generate this type of descriptor is illustrated in detail. Theoretical analyses and experimental verifications show that the proposed method has better performance than those of other methods included in the experiments in complicated situations. We adopt the method of angle traversal instead of assigning a dominant orientation and consider all the feature points as a whole to make the descriptor independent of the gray levels around the feature points. Therefore, the descriptor is insensitive to changes in illumination. Moreover, the designed descriptor maintains high robustness when inversion of gray levels occurs in remote-sensing images.

### Conflicts of interest

The authors declare no conflicts of interest.

### Acknowledgement

This work was supported by the Science and Technology Development Program of Jilin under Grant No. 20170204029GX.

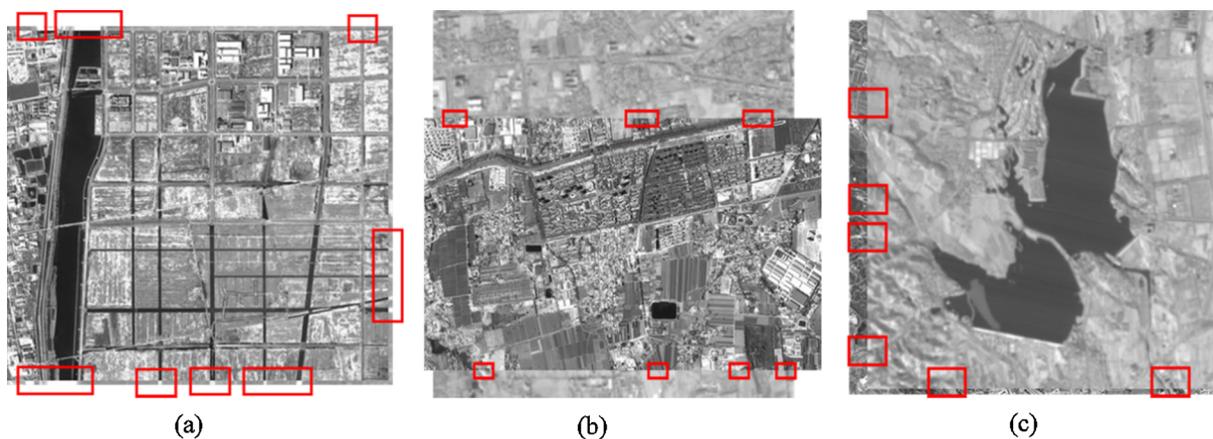


Fig. 18. The overlapping images of the original visible-light images and the far-infrared images. (a) The overlapping image in the fourth group. (b) The overlapping image in the fifth group. (c) The overlapping image in the sixth group.

## References

- [1] B. Zitová, J. Flusser, Image registration methods: a survey, *Image Vis. Comput.* 21 (11) (2003) 977–1000.
- [2] M. Ehlers, Multisensor image fusion techniques in remote sensing, *Isprs J. Photogram. Remote Sens.* 46 (1) (1991) 19–30.
- [3] C. Shi, Q. Miao, P. Xu, A novel algorithm of remote sensing image fusion based on Shearlets and PCNN, *Neurocomputing* 117 (14) (2013) 47–53.
- [4] M. Ghahremani, H. Ghassemian, Remote sensing image fusion using ripplet transform and compressed sensing, *IEEE Geosci. Remote Sens. Lett.* 12 (3) (2014) 502–506.
- [5] V. Alberga, Similarity measures of remotely sensed multi-sensor images for change detection applications, *Remote Sens.* 1 (3) (2009) 122–143.
- [6] L. Jia, M. Li, Y. Wu, P. Zhang, G. Liu, H. Chen, L. An, SAR image change detection based on iterative label-information composite kernel supervised by anisotropic texture, *IEEE Trans. Geosci. Remote Sens.* 53 (7) (2015) 3960–3973.
- [7] M. Miura, S. Sakai, S. Aoyama, J. Ishii, K. Ito, T. Aoki, High-accuracy image matching using phase-only correlation and its application, *Sice Conference*, 2012, pp. 307–312.
- [8] G. Tzimiropoulos, V. Argyriou, T. Stathaki, Subpixel registration with gradient correlation, *IEEE Trans. Image Process. A Publ. of the IEEE Signal Process. Soc.* 20 (6) (2011) 1761.
- [9] J. Inglada, V. Muron, D. Pichard, T. Feuvrier, Analysis of artifacts in subpixel remote sensing image registration, *IEEE Trans. Geosci. Remote Sens.* 45 (1) (2007) 254–264.
- [10] C. Harris, A combined corner and edge detector, *Proc. Alvey Vision Conf.* 1988 (3) (1988) 147–151.
- [11] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vision* 60 (2) (2004) 91–110.
- [12] H. Bay, T. Tuytelaars, L.V. Gool, SURF: speeded up robust features, *European Conference Computer Vision*, 2006, pp. 404–417.
- [13] C.Y. Lee, H.J. Wang, J.H. Lai, Y.C. Chang, C.S. Huang, Automatic marker-free longitudinal infrared image registration by shape context based matching and competitive winner-guided optimal corresponding, *Sci. Rep.* 7 (2017) 39834.
- [14] J. Jiang, S. Cao, G. Zhang, Y. Yuan, Shape registration for remote-sensing images with background variation, *Int. J. Remote Sens.* 34 (15) (2013) 5265–5281.
- [15] C. Zhao, H. Zhao, J. Lv, S. Sun, B. Li, Multimodal image matching based on multimodality robust line segment descriptor, *Neurocomputing* 177(C) (2016) 290–303.
- [16] Y. Li, R.L. Stevenson, Multimodal image registration with line segments by selective search, *IEEE Trans. Cybern.* 47 (5) (2017) 1285–1298.
- [17] Z. Song, S. Zhou, J. Guan, A novel image registration algorithm for remote sensing under affine transformation, *IEEE Trans. Geosci. Remote Sens.* 52 (8) (2014) 4895–4912.
- [18] H.S. Stone, M.T. Orchard, E.C. Chang, S.A. Martucci, A fast direct Fourier-based algorithm for subpixel registration of images, *Geosci. Remote Sens. IEEE Trans.* on 39 (10) (2001) 2235–2243.
- [19] J. Thornton, M. Savvides, B.V.K.V. Kumar, A Bayesian approach to deformed pattern matching of iris images, *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (4) (2007) 596–606.
- [20] H. Foroosh, J.B. Zerubia, M. Berthod, Extension of phase correlation to subpixel registration, *IEEE Trans. Image Process. A Publ. IEEE Signal Process. Society* 11 (3) (2002) 188–200.
- [21] M. Gong, S. Zhao, L. Jiao, D. Tian, S. Wang, A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information, *IEEE Trans. Geosci. Remote Sens.* 52 (7) (2014) 4328–4338.
- [22] M. Brown, D.G. Lowe, Automatic panoramic image stitching using invariant features, *IJCV* (2007) 59–73.
- [23] J. Luo, O. Gwun, A comparison of SIFT, PCA-SIFT and SURF, *Int. J. Image Process.* 3 (4) (2013) 143–152.
- [24] M. Calonder, V. Lepetit, C. Strecha, P. Fua, BRIEF: binary robust independent elementary features, *European Conference on Computer Vision*, 2010, pp. 778–792.
- [25] E. Rosten, T. Drummond, Machine learning for high-speed corner detection, *European Conference on Computer Vision*, 2006, pp. 430–443.
- [26] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, ORB: An efficient alternative to SIFT or SURF, *IEEE International Conference on Computer Vision*, 2012, pp. 2564–2571.
- [27] S. Leutenegger, M. Chli, R.Y. Siegwart, BRISK: Binary Robust invariant scalable keypoints, *IEEE International Conference on Computer Vision*, 2012, pp. 2548–2555.
- [28] L. Yan, D. Roy, H. Zhang, J. Li, H. Huang, An automated approach for sub-pixel registration of landsat-8 operational land imager (OLI) and sentinel-2 multi spectral instrument (MSI) imagery, *Remote Sens.* 8 (6) (2016) 520.
- [29] C. Lyu, J. Jiang, Remote sensing image registration with line segments and their intersections, *Remote Sens.* 9 (5) (2017) 439.
- [30] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, K.R. Mullers, Fisher discriminant analysis with kernels, *Neural Networks for Signal Processing IX*, 1999, Proceedings of the 1999 IEEE Signal Processing Society Workshop, 2002, pp. 41–48.