Research Article

Tao Shu*

Multi-sensor remote sensing image alignment based on fast algorithms

Abstract: Remote sensing image technology to the ground has important guiding significance in disaster assessment and emergency rescue deployment. In order to realize the fast automatic registration of multi-sensor remote sensing images, the remote sensing image block registration idea is introduced, and the image reconstruction is processed by using the conjugate gradient descent (CGD) method. The scale-invariant feature transformation (SIFT) algorithm is improved and optimized by combining the function-fitting method. By this way, it can improve the registration accuracy and efficiency of multi-sensor remote sensing images. The results show that the average peak signal-to-noise ratio of the image processed by the CGD method is 25.428. The average root mean square value is 17.442. The average image processing time is 6.093 s. These indicators are better than the passive filter algorithm and the gradient descent method. The average accuracy of image registration of the improved SIFT registration method is 96.37%, and the average image registration time is 2.14 s. These indicators are significantly better than the traditional SIFT algorithm and speeded-up robust features algorithm. It is proved that the improved SIFT registration method can effectively improve the accuracy and operation efficiency of multi-sensor remote sensing image registration methods. The improved SIFT registration method effectively solves the problems of low accuracy and long time consumption of traditional multi-sensor remote sensing image fast registration methods. While maintaining high registration accuracy, it improves the image registration speed and provides technical support for a rapid disaster assessment after major disasters such as earthquakes and floods. And it has an important value for the development of the efficient post-disaster rescue deployment.

Keywords: remote sensing, image registration, SIFT, function fitting, conjugate gradient decreases, sensor, Hessian matrix

1 Introduction

Remote sensing earth observation technology can detect and analyse changes in the ground environment in real time. After major disasters such as earthquakes, floods, and forest fires, remote sensing earth observation technology provides a stable, fast, and safe information acquisition method for disaster assessment and rapid emergency response after disasters [1,2]. Using remote sensing to ground monitoring technology to compare and analyse ground remote sensing images before and after the occurrence of a disaster can provide assistance for a rapid acquisition of disaster situations and criticality assessment, and help relevant departments quickly launch emergency response mechanisms based on the specific situation of the disaster, deploy personnel to rescue, and other post-disaster work. Multi-sensor remote sensing image registration is an important foundation for quickly extracting disaster information. The accuracy and speed of remote sensing image registration directly affect the efficiency of taking emergency measures after a disaster. It is required to

---

* **Corresponding author: Tao Shu,** Information Technology Center, Sichuan Vocational and Technical College, Suining 629000, China, e-mail: slyst5201@163.com

strengthen the analysis of disaster information and improve the efficiency of disaster assessment as much as possible. However, the imaging mechanism of different sensors is different, and the image shooting time and angle are also different. The remote sensing images obtained by different sensors have different resolution and information characteristics. Therefore, it is necessary to process the remote sensing images obtained by different sensors and match them with the ground remote sensing images of the disaster-affected areas before the disaster, so as to achieve a rapid acquisition and evaluation of disaster information [3,4]. At present, the registration of large-scale remote sensing images from different sensors is inefficient. Therefore, the research takes large-scale ground remote sensing images as the research object, focusing on the registration of multi-sensor remote sensing images after disasters, in order to achieve a rapid automatic registration of remote sensing images and provide reference for disaster assessment and emergency rescue.

In order to improve the registration accuracy and efficiency of remote sensing images from different sensors, the image block registration method is proposed, and the conjugate gradient descent (CGD) method is used for image reconstruction. The scale-invariant feature transform (SIFT) algorithm and the function-fitting method are combined to extract and match image feature points, hoping to provide help and reference for the rapid disaster assessment. The main contribution of the research is to introduce the idea of segmentation for image registration and combine the CGD method for image reconstruction, effectively solving the problem of low efficiency in large-scale remote sensing image registration and improving the efficiency of automatic image registration. Second, an improved fast automatic image registration technology was proposed. The SIFT algorithm was improved and optimized using the function-fitting method. Combining the image block registration idea and the CGD method, fast and large-scale ground remote sensing image registration was effectively achieved.

2 Related works

Fast algorithms are a pillar of signal processing. With the continuous development of information technology, research on fast algorithms is increasing. Bowman and Ghoggali [5] proposed an algorithm for computing one-dimensional partial fast Fourier changes, in which the null-frequency domain is divided into rectangular and trapezoidal sub-domains. The fast algorithm is developed based on the subdomains. And it is demonstrated computationally that the partial Fourier transform in the algorithm can be reduced to convolution for fast computation. To solve the problem of difficulty in determining the tile size of Winograd-based convolution, they designed a tile fusion method that can derive the optimal tile size in each convolution layer. It was demonstrated that this method could accelerate the convolution inference process [6]. It was experimentally demonstrated that it could improve the performance of Winograd convolution by a factor of 1.89. Fan et al. [7] investigated the integral formulation and fast algorithm for the steady-state radiative transfer equation. Fourier transform and the spherical harmonic transform coefficients were used to reduce the number of steps in the computational process. And the number of degrees of freedom in the computational process was reduced. By this way, the Fourier transform and spherical harmonic transform coefficients were obtained. The effectiveness and significant computational efficiency of the algorithm proposed in this study were demonstrated in numerical simulations. Li et al. [8] argued that in the modern information society where the amount of data was increasing, the widespread use of convolutional neural networks has led to a gradual decrease in their efficiency in the computation process. In order to solve these problems, a fast convolutional algorithm was proposed, which was able to reduce both the number of computations and data accesses in the computation process. Wang et al. [9] proposed a fast algorithm based on the non-uniform Fourier transform and used it to compute non-uniform sparse aperture near-fields. Simulation experiments showed that the algorithm had the effect of accelerating the near-field calculation of non-uniform sparse aperture. It is proved that the algorithm had a significant universality.

With the continuous development of modern technology, remote sensing technology had become a key research object in technological development. Jayanthi and Vennila [10] proposed a classification method based on a novel deep neural network classifier in order to obtain the same remote sensing view through the

accurate classification. In this method, linear iterative clustering combined with neural network was used to process satellite images. It was shown that the method could achieve dynamic alignment of remote sensing images. Wu et al. [11] proposed a fast robustness-based sub-pixel remote sensing image alignment method in order to solve the remote sensing image alignment problem under uncertainties. In the experimental results, it was shown that the method proposed in this study was able to obtain more anti-interference matches compared to other methods. And the computational effort in the image matching process was smaller. Jia et al. [12] proposed a fast robustness-based sub-pixel remote sensing image alignment method in order to obtain effective image alignment in wide. In order to obtain effective image dense matching in wide baseline images, they proposed a matching method based on proportional triangulation, which eliminated proportional points by calculating the positional similarity of the triangulation points corresponding to the two images. It showed the effectiveness and accuracy of the method in matching experiments, which is of great value for remote sensing image dense alignment. Paul et al. [13] applied the SIFT algorithm to remote sensing image alignment. The SIFT algorithm was found to be able to reduce the effect of speckle noise in SAR image alignment. It was experimentally demonstrated that the proposed method could increase the positional accuracy in alignment. Ma et al. [14] proposed a new remote sensing image alignment method, which eliminated the effect of adjacent points by phase consistency feature detection. It demonstrated that the method had better alignment performance compared with the current main alignment methods.

In summary, the study of remote sensing image alignment had become the focus of current research. And it was also known that the alignment efficiency in remote sensing image alignment has always been an important concern for existing scholars. Fast algorithms are efficient in the computational process, but they are less used in remote sensing image alignment. Therefore, in order to improve the accuracy and efficiency of multi-sensor remote sensing image alignment, we try to apply the fast algorithm to multi-sensor remote sensing image alignment. It is hoped to provide theoretical support for the development of remote sensing technology in China.

3 Fast alignment of multi-sensor remote sensing images based on improved SIFT

3.1 CGD-based image reconstruction and image chunking

CGD has obvious advantages in solving large-scale unconstrained optimization problems. It has the advantages of rapid convergence, small data storage requirements, and secondary termination. In order to solve the problems of inefficient operation and low accuracy of traditional multi-sensor remote sensing image alignment methods, the research is based on the idea of image chunking alignment. The research uses CGD to pre-process multi-sensor remote sensing images, and it is combined with SIFT algorithm, nearest neighbour method, and function-fitting method to extract and match the image feature points. The process of multi-sensor image alignment is shown in Figure 1.
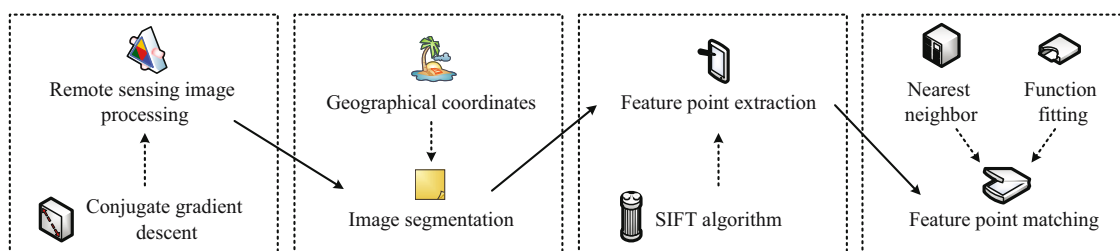

Figure 1: Multi-sensor image registration process.

In order to improve the accuracy of remote sensing image alignment, pre-processing of remote sensing images is required to enhance the information accuracy of remote sensing images. This study makes use of the information complementarity between multiple remote sensing images of the same region to achieve image recovery and reconstruction. And it uses the CGD method, which has the advantage of convergence speed. By this way, it can reconstruct the images and improve the efficiency of real-time image processing while ensuring image quality. The relationship between the pixel values in the lower-resolution image and the corresponding high-resolution image is shown in the following equation:

$$b_m = \sum_{r=1}^{N} H_{m,r} x_r + \eta_m, \tag{1}$$

where b_m is the element of m in the low-resolution image, x_r is the element of r in the high-resolution image, $H_{m,r}$ is the weighting effect of the high-resolution pixel of r on b_m, and η_{m^*} is the effect of additional noise on the value of the pixel of m, and $m = 1, 2, \cdots, M$. A regularized cost function is introduced to estimate the high-resolution image x and the regularized cost function is shown in the following equation:

$$C(x) = \frac{1}{2} \sum_{m=1}^{pM} \left(b_m - \sum_{r=1}^{N} H_{m,r} x_r \right) + \frac{\lambda}{2} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} a_{i,j} x_j \right)^2, \tag{2}$$

where λ is the weighted control parameter and $a_{i,j}$ is the regularization parameter. The conjugate gradient iterative meter for the image is shown in the following equation:

$$\hat{x}_k^{n+1} = \hat{x}_k^{n} + \varepsilon^n d^k(\hat{x}^n), \tag{3}$$

where n is the number of iterations, $n = 0, 1, 2, \cdots, N$. k is the variable numbers, $k = 0, 1, 2, \cdots, N$. $d^k(\hat{x}^n)$ is the conjugate gradient term. ε^n is the step size of the nth iteration. The step size ε^n is obtained by minimizing the cost function as shown in the following equation:

$$\varepsilon^n = -\frac{\sum_{m=1}^{pM} \phi_m (\sum_{r=1}^{N} H_{m,r} \hat{x}_r^n - b_m)}{\sum_{m=1}^{pM} \phi_m^2 + \lambda \sum_{i=1}^{N} \overline{d_i^2}} - \lambda \frac{\sum_{i=1}^{N} d_i (\sum_{j=1}^{N} a_{i,j} \hat{x}_r^n)}{\sum_{m=1}^{pM} \phi_m^2 + \lambda \sum_{i=1}^{N} \overline{d_i^2}}, \tag{4}$$

where $\phi_m = \sum_{r=1}^{N} H_{m,r} d_r(\hat{x}_r^n)$, $d_i = \sum_{j=1}^{N} a_{i,j} d_j(\hat{x}_r^n)$, d_i represents the conjugate gradient vector when the step size is i, and d_j represents the conjugate gradient vector when the step size is j. Feature extraction of large-format remote sensing images requires a large amount of memory and computation and takes a lot of time. Therefore, this study uses geographic coordinate constraints as the basis for feature extraction and matching of the target image in blocks. And it matches the reference image based on the projection coordinates of the segmented image blocks to achieve feature matching for each image block. The mapping relationship between remote sensing images and geographical coordinates is shown in the following equation:

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} + \begin{bmatrix} G_1 & G_2 \\ G_4 & G_5 \end{bmatrix} \begin{bmatrix} I_i \\ J_i \end{bmatrix}, \tag{5}$$

where (I_i, J_i) is the image coordinate of the i point in the image, (X_i, Y_i) is the corresponding projection coordinate, and (X_0, Y_0) is the geographical coordinate of the image. $G_1, G_2, G_4,$ and G_5 are the transformation parameters. The image chunking idea is shown in Figure 2. The target image is divided into $M \times N$ image chunks, and the image chunks are numbered and sorted. The four corners of the image chunks are marked in a clockwise order, and the image coordinates of the upper-left and lower-right corners of the image chunks are recorded. The image coordinates are converted to projection coordinates based on the geographic coordinates of the target image, combined with the image element size G_1 and its negative value G_5. According to the projection coordinates, the reference image block corresponding to the target image is determined. And the chunked image features of the target image are matched to improve the efficiency of the target image alignment.
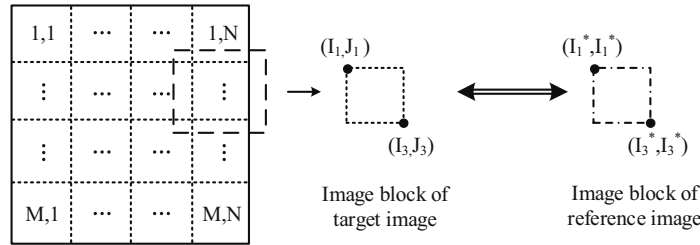
Figure 2: Image block processing.

3.2 SIFT-based detection of extreme value points

Between multiple remote sensing images of the same region and geometrically aligning the multi-sensor remote sensing images, there is a mapping relationship. By solving the mapping relationship, remote sensing image alignment can help subsequent image feature analysis and comparison of regional information changes [15,16]. The SIFT algorithm takes point features as the starting point for image alignment and has advantages in obtaining stable local features. Therefore, the SIFT algorithm is used for the fast alignment of multi-sensor remote sensing images. The feature matching process of the SIFT algorithm is shown in Figure 3.
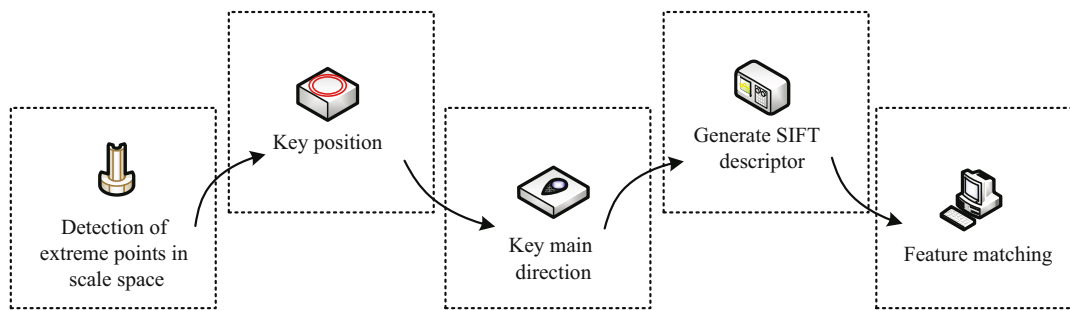

Figure 3: Feature matching process of SIFT algorithm.

The SIFT algorithm uses a Gaussian pyramid to construct different scale spaces and a Gaussian kernel function to process the target image. So the resolution of the target image of different sizes is the same, and the Gaussian-scale space of the target image is $L(x, y, \sigma)$. The construction function of the target image is shown in the following equation:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \tag{6}$$

where $G(x, y, \sigma)$ is the 2D Gaussian kernel function, $*$ represents the convolution operation, and $I(x, y)$ is the initial image before the operation. The two-dimensional Gaussian kernel function $G(x, y, \sigma)$ is defined in the following equation:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2}, \tag{7}$$

where σ is the standard deviation of the Gaussian distribution, indicating the degree of blurring of the initial image. The size of σ is positively correlated with the degree of blurring and the corresponding scale of the image. The smaller the scale of the image, the better it reflects the detailed features of the image. However, the larger the scale of the image, the better it reflects the overall general characteristics of the image. Therefore, it is necessary to determine an appropriate σ value to construct the Gaussian scale space of the target image [17]. The SIFT algorithm detects the feature points common to different scales of images by constructing a Gaussian difference scale space and convolves the difference between the target image and the adjacent-scale space

function. The SIFT algorithm is based on the Gaussian differential-scale space of the image and detects the extreme points of the image at different scales [18]. The definition of the Gaussian difference scale space of an image is shown in the following equation:

$$
\begin{aligned}
D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\
&= L(x, y, k\sigma) - L(x, y, \sigma),
\end{aligned}
\tag{8}
$$

where k is the neighbouring-scale space factor scale factor. By comparing the two layers of images adjacent to each Gaussian difference space, local extreme value points in the Gaussian difference scale space are obtained. To find the extreme value points in the space with the largest and smallest grey pixel values, the grey pixel points in the image are compared with all neighbouring pixel points.

3.3 Extreme point filtering based on Hessian matrix

The polar points obtained from the discrete space have error problems and are not necessarily the true polar points of the image. So to improve the accuracy of the image alignment results, the polar points obtained need to be filtered to eliminate the edge response feature points with low contrast and poor stability [19,20]. The Taylor formula is used to locate the location of the feature points and determine the scale space where the feature points are located. The Taylor expansion of the scale space Gaussian difference function for the pixel location is shown in the following equation:

$$
D(X) = D(X_0) + \frac{\partial D^T}{\partial X}(X - X_0) + \frac{1}{2}(X - X_0)^T \frac{\partial^2 D}{\partial X^2}(X - X_0),
\tag{9}
$$

where X is the accurate location and scale information of the feature points, and $X = x, y, \sigma$. X_0 is the location and scale information of the feature points obtained from the Gaussian difference scale space, that is, the interpolation centre. T is the matrix transposition. Use equation (9) to derive X and make the result 0 to obtain the offset of extreme point. The offset function of extreme point $\hat{X}$ relative to X_0 is shown in the following formula:

$$
\hat{X} = \frac{\partial^2 D^{-1}}{\partial X^2}\frac{\partial D}{\partial X},
\tag{10}
$$

where $X = \hat{X} + X_0$. Then, the offset is brought into equation (9). It can give the response value of the feature point at the new fitted position as shown in the following equation:

$$
D(\hat{X}) = D(X_0) + \frac{1}{2}\frac{\partial D^T}{\partial X}\hat{X}.
\tag{11}
$$

Set the minimum response threshold for the low contrast feature points that need to be removed, and filter all feature points. If the response value of the feature point $\|D(\hat{X})$ is less than the threshold, the feature point is judged to be a low-contrast point with poor stability and vulnerable to noise and needs to be eliminated.

The strong edge response is an important feature of Gaussian difference images, and the extreme points are detected by the Gaussian difference scale space. It suffers from the problem of small principal curvature in the edge direction and large principal curvature in the edge gradient direction. Therefore, this study uses the Hessian matrix to reject the edge response feature points and analyses the principal curvature of the Gaussian difference function of the feature points by the eigenvalues of the Hessian matrix. The Hessian matrix function of the feature points is shown in the following equation:

$$
H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix},
\tag{12}
$$

where D_{xx} and D_{yy} represent the second derivatives of the feature points in the direction of x and y in the Gaussian differential scale space, respectively. D_{xy} and D_{yx} represent the second-order mixed partial

derivatives of the feature points in the Gaussian differential scale space, respectively. The trace and determinant of the matrix are shown in the following equation:

$$\begin{cases} \text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta \\ \text{Det}(H) = D_{xx}D_{yy} - D_{xy}^2 = \alpha\beta \end{cases},$$

(13)

where α and β represent the maximum and minimum eigenvalues of the matrix, respectively. The magnitude of the principal curvature of the eigenpoints is obtained by calculating the ratio of the eigenvalues of the Hessian matrix. Let the eigenvalue ratio be ς, and the constraints on the eigenvalue ratio are shown in the following equation:

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(\varsigma + 1)^2}{\varsigma}.$$

(14)

There is a positive correlation with the principal curvature of the feature point. When the feature point satisfies equation (14), the feature point is judged to be an edge response point that needs to be rejected.

3.4 Feature matching based on nearest neighbour and function fitting

The SIFT algorithm is used to extract the feature points in the target image and assign the base direction for feature point matching. Combining the contribution relationship between the feature points and the domain space pixel points, the histogram of the gradient direction of the feature points to the domain space pixels is constructed to achieve the main direction matching of the feature points. The gradient direction $\theta(x, y)$ and magnitude $m(x, y)$ of the domain pixels of the feature point are shown in the following equation:

$$\begin{cases} \theta(x, y) = \arctan\left(\dfrac{L(x, y + 1) - L(x, y - 1)}{L(x + 1, y) - L(x - 1, y)}\right) \\ m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \end{cases},$$

(15)

where L is the Gaussian differential scale space of the feature points. The direction information of feature points is combined with coordinate information and scale information to form the descriptors of feature points. It can ensure the correct matching of feature points in the same scene and the differentiation of feature points in different scenes. In order to ensure that the descriptors of feature points remain stable to changes such as rotation and scaling of images, first, it needs to make the domain gradient direction of feature points consistent with the main direction of feature points. The domain position with feature points is selected as the centre. And the sub-region with the side length of 3σ is divided. The square domain is divided into 4×4 sub-regions equally. Let the main direction of the feature point be θ, then the adjusted field pixel point coordinates are shown in the following equation:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

(16)

where (x, y) and (x', y') represent the coordinates of the pixel points before and after adjustment, respectively. The nearest neighbour method is used to match the feature points and measure whether there is a pairwise relationship between two feature points. By calculating the Euclidean distance between the reference image feature point and the target image feature point, the nearest neighbour method searches for the point in the image with the closest Euclidean distance to the feature point, which is the pairing point of the feature point. Let the reference image feature point be P, the feature point in the target image be Q, and the Euclidean distance between the two points be $D(P, Q)$. The calculation function is shown in the following equation:

$$D(P, Q) = \sqrt{\sum_{n=1}^{128} (P_n - Q_n)^2},$$

(17)

where P_n and Q_n denote the nth dimension of the P and Q descriptors, respectively. In order to improve the accuracy of feature point matching, the function-fitting method is investigated to further filter the matching points and eliminate the incorrect matching points among them. The operation flow of the function-fitting method is shown in Figure 4. A function model is constructed based on the feature information of the reference image and the feature points of the target image. And the function model is solved by combining iterative least-squares fitting method to obtain an accurate function model of the target image. The error between the feature points and the function model of the target image is calculated, and an error threshold is set. When the calculation error of the matching point is higher than the threshold, the matching point is rejected.
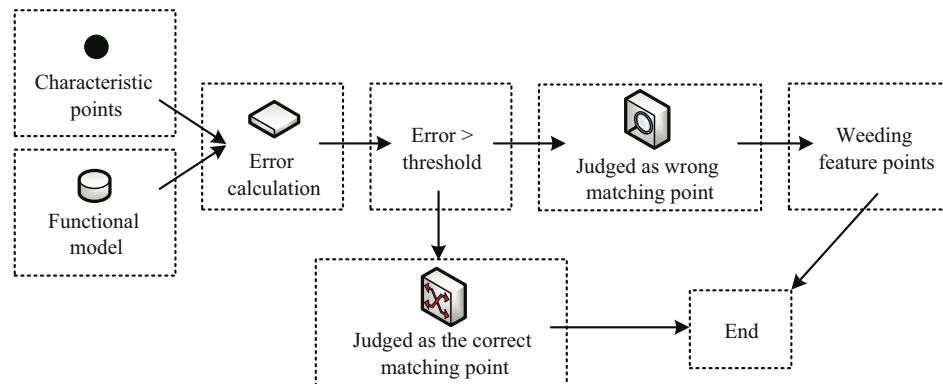

Figure 4: Operation process of the function-fitting method.

A cubic polynomial function model of the target image is constructed. And the function model is solved using least squares to make the data satisfy the function parameters as much as possible. It is achieved by minimizing the error-squared and by matching the best function. Let the function to be fitted be $f(x)$ and the actual data be x_i. The least-squares fitted function is shown in the following equation:

$$A = \min \sum_{i=1}^{r} [f(x_i) - x_i]^2, \tag{18}$$

where r is the feature point of the target image. The accurate function model of the target image is solved through repeated iterations. The error between the function model and the feature points is calculated. The two matching points with the largest error for each iteration are eliminated. And the error of all points is guided to be less than the threshold to obtain the accurate function model. The error between all matching points and functions is calculated. The error threshold σ is set as the mean value of all errors. And the hard threshold shrinkage method is used to compare the error value and the threshold value. The hard threshold shrinkage method function is shown in the following equation:

$$\text{erf} = \begin{cases} \text{erf}(i, j), & |\text{erf}(i, j)| \geq \sigma \\ 0, & |\text{erf}(i, j)| < \sigma. \end{cases} \tag{19}$$

4 Effect of multi-sensor remote sensing image alignment

4.1 Image reconstruction and feature matching effect analysis

In order to verify the optimality of the CGD method used in this study for image reconstruction, the CGD algorithm was compared with the passive filter algorithm and the steepest descent (SD) method for the application. The peak

signal-to-noise ratio (PSNR) and root mean square (RMS) metrics were used to evaluate the image reconstruction effects of the three algorithms, and the comparison of the three algorithms is shown in Figure 5.
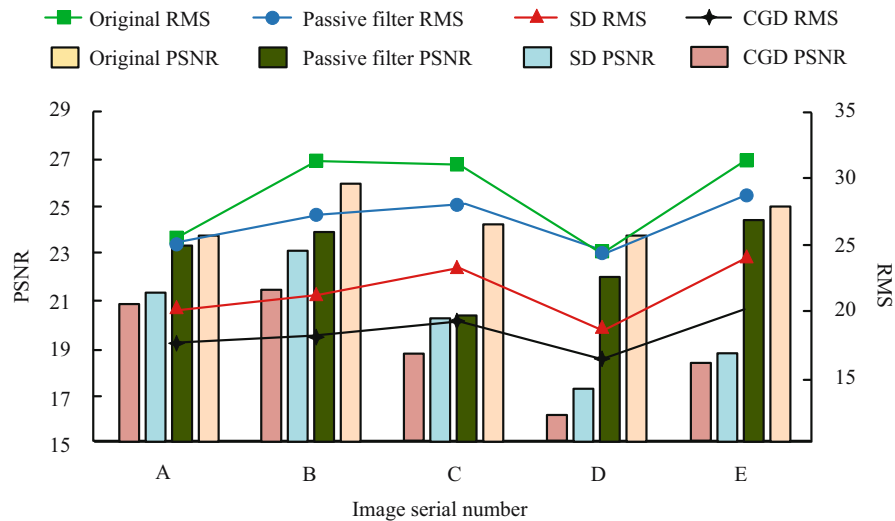

Figure 5: Comparison of image reconstruction effects of three algorithms.

As can be seen in Figure 5, the PSNR values of remotely sensed images with image processing by the regularized CGD method are significantly higher than those of images processed by the passive filter algorithm and the gradient descent method. The average PSNR value of the CGD method is 25.428, which is higher than that of the passive filter algorithm (21.268) and the SD method (24.356). While the average RMS value of the images processed by the CGD method is 17.442, which is significantly lower than that of the passive filter algorithm and the gradient descent method. To explore the difference in image reconstruction efficiency of the three algorithms, the processing time of the three algorithms for ten images was compared and analysed. The image reconstruction time consumption of the three algorithms is shown in Table 1.

Table 1: Reliability analysis of college students' values scale

Image number	Passive filter (s)	SD (s)	CGD (s)
1	7.545	6.485	6.586
2	7.596	6.754	6.841
3	6.587	6.255	5.067
4	7.698	6.567	6.345
5	6.265	6.261	6.052
6	7.066	6.712	6.593
7	7.485	6.263	6.841
8	6.281	6.485	5.441
9	7.062	6.466	5.046
10	6.592	6.257	6.155
11	6.891	6.304	6.702
12	7.034	6.724	6.138
13	7.358	6.105	5.681
14	6.942	6.381	6.104
15	7.423	6.158	6.294

As can be seen from Table 1, the average time taken by the passive filter method for processing ten images is 7.055 s. While the average time taken by the SD method on image processing is 6.412 s, the average image processing time for the CGD method is 6.126 s, which is significantly lower than the other two algorithms. It is proved that the regularized CGD method can improve image processing while ensuring higher image processing quality speed and efficiency.

This study verified the matching effect of the image feature points and analysed the performance of the function-fitting method on the rejection of false match points. In this study, a comparative analysis of the function fitting before and after the rejection of false match points was performed, as shown in Figure 6. Figure 6a shows the fit after the initial matching by the nearest neighbour method, and Figure 6b shows the fit after the false match points are removed by the function-fitting method.
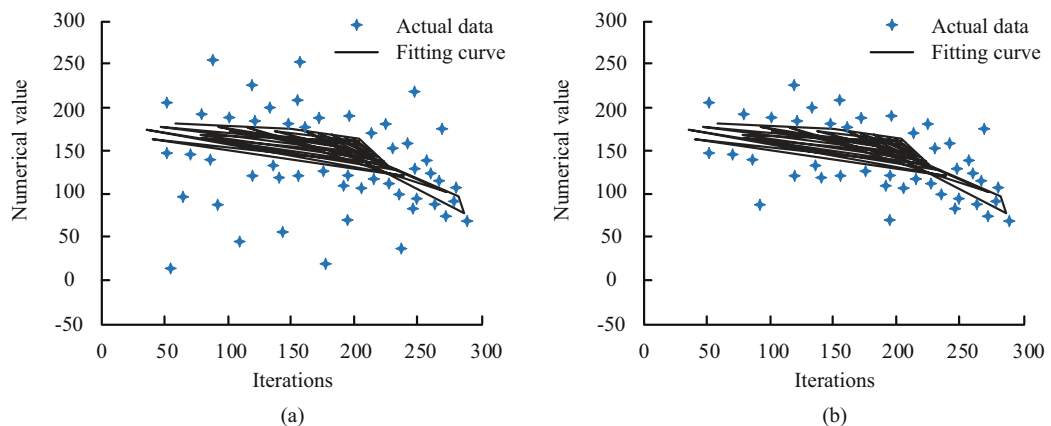

Figure 6: Function fitting before and after removing the wrong matching points. (a) Fitting results before removing incorrect matching points and (b) Fitting results after removing incorrect matching points.

From Figure 6, the initial matching using the nearest neighbour method had certain matching errors and the data fit was poor. After further screening of the matching points using the function-fitting method, the wrong feature points that were farther away from the fitted curve are eliminated. It was proved that the function-fitting method could effectively eliminate the wrong matching points in the feature matching process and improve the accuracy of image alignment.

4.2 Comparative experiment and application effect analysis of image registration

This study verified the effectiveness and efficiency of the multi-sensor remote sensing image alignment method based on the improved SIFT algorithm. In this study, it used the improved SIFT alignment method, the traditional SIFT algorithm, and the speeded-up robust features (SURF) algorithm to align 70 pairs of remote sensing images. And the image alignment accuracy of the three algorithms is shown in Figure 7. The image alignment accuracy of the three algorithms is shown in Figure 7.

From Figure 7, it can be seen that the image alignment accuracy of the traditional SIFT algorithm is the lowest for 70 pairs of remote sensing images, with an average accuracy of 80.49%. And the image alignment accuracy of the SURF algorithm is 84.95%. However, the average accuracy of the improved SIFT alignment method is 96.37%, which is significantly higher than that of the traditional SIFT algorithm and the SURF algorithm. The accuracy of image alignment fluctuated the least, which proved that the improved SIFT alignment method was more robust. Four pairs of images were selected from the 70 pairs for the combined analysis of accuracy and recall, and the recall and 1-precision curves of the four pairs are shown in Figure 8.
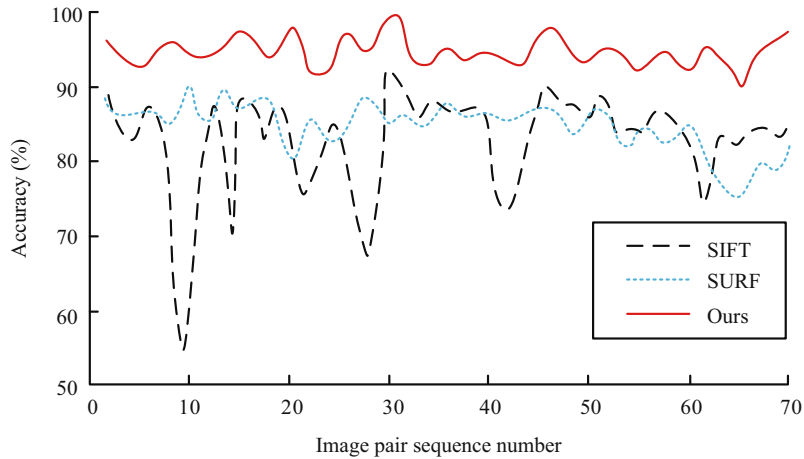
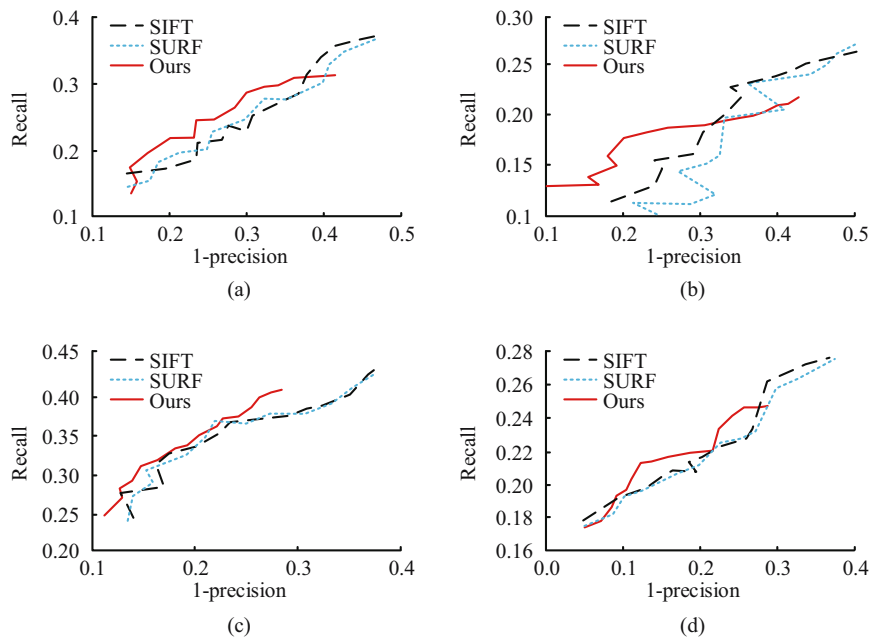**Figure 7:** Image registration accuracy of three algorithms.

Figure 8: Recall and 1-precision curves of four pairs of images. (a) Image pair a, (b) image pair b, (c) image pair c, and (d) image pair d.

As can be seen from Figure 8, the overall trend is that the recall value increases with the 1-precision value, and the accuracy and recall are inversely correlated. The accuracy and recall of the improved SIFT alignment method outperformed the traditional SIFT algorithm and the SURF algorithm. It is proved that the improved SIFT alignment method has advantages in the alignment performance of remote sensing images. The image alignment time of the three algorithms was compared and analysed for 70 pairs of remote sensing images, and the image alignment operation efficiency of the three algorithms was analysed. The image alignment time consumption of the three algorithms is shown in Figure 9.

As can be seen from Figure 9, the traditional SIFT algorithm takes the longest time in image alignment, with an average time of 6.89 s, while the average time of SURF algorithm is 4.08 s. The average time of the improved SIFT alignment method is 2.14 s for 70 pairs of images, which is 4.75 and 1.94 s lower than that of the traditional SIFT algorithm and SURF algorithm. This demonstrates that the improved SIFT alignment method can effectively improve the efficiency of remote sensing image alignment and shorten the image alignment time. The improved SIFT alignment method improves the image alignment speed by the method of image
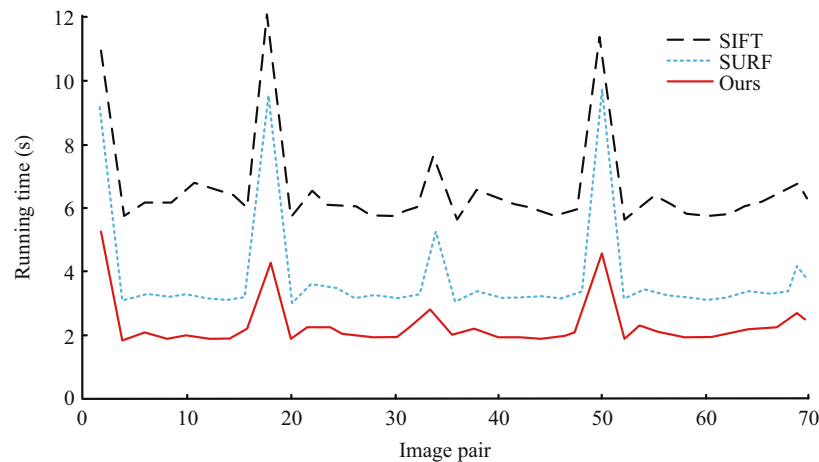
Figure 9: Time-consuming image registration of three algorithms.

chunking and introduces an efficient feature point matching method. By this way, it can effectively solve the problem of inefficient operation of the traditional SIFT algorithm.

5 Conclusion

In order to improve the accuracy and efficiency of multi-sensor image alignment and achieve rapid acquisition and monitoring of disaster situations, this study proposes a new method. It uses the CGD method for image reconstruction processing based on the idea of image chunking alignment. The improved SIFT algorithm is used to extract and match the feature points of remote sensing images. And to achieve a rapid automatic alignment of multi-sensor remote sensing images, it combines the function-fitting method to further propose the error matching points. The results show that the average PSNR value of the image under the CGD method processing is 25.428, the average RMS value is 17.442, and the average image processing time is 6.093 s. The image reconstruction effect and efficiency are better than those of the passive filter algorithm and the gradient descent method. The introduction of the function-fitting method in the SIFT algorithm can effectively eliminate the wrong matching points in the feature point matching process and improve the image feature point matching accuracy. The average accuracy of the improved SIFT alignment method is 96.37%, and the average image alignment time is 2.14 s. These are significantly better than the traditional SIFT algorithm and SURF algorithm in terms of image alignment accuracy and efficiency. This method has stronger robustness. The research takes point features as the entry point to register remote sensing images. In future, other effective feature elements such as area features and line features can be introduced for further analysis. It can be combined with point feature analysis to further improve the accuracy and speed of multi-sensor image registration.

References

[1]   He H, Chen T, Chen M, Li D, Cheng P. Remote sensing image super-resolution using deep-shallow cascaded convolutional neural networks. Sens Rev. 2019;39(5):629–35.

[2]   Alonso-Gonzalez L, Arboleya A, Las-Heras F. Astursat: A software tool to encourage the interest of students in remote sensing and image processing. Int J Eng Educ. 2019;35(3):912–24.

[3]   Yusuke S, Daisuke B, Jessie BJ, Yatagai T. Spherical-harmonic-transform-based fast calculation algorithm for spherical computer-generated hologram considering occlusion culling. Appl Opt. 2018;57(23):6781–7.

[4]   Daliakopoulos IN, Tsanis IK. A SIFT-based DEM extraction approach using GEOEYE-1 satellite stereo pairs. Sensors. 2019;19(5):1123–40.

[5]   Bowman JC, Ghoggali Z. The partial fast Fourier transform. J Sci Comput. 2018;76(3):1578–93.

[6]   Ji Z, Zhang X, Wei Z, Li J, Wei J. A tile-fusion method for accelerating winograd convolutions. Neurocomputing. 10 Oct 2021;460:9–19.

[7]   Fan Y, An J, Ying L. Fast algorithms for integral formulations of steady-state radiative transfer equation. J Comput Phys. 2018;380:191–211.

[8]   Li WJ, Ruan SJ, Yang DS. Implementation of energy-efficient fast convolution algorithm for deep convolutional neural networks based on FPGA. Electron Lett. 2020;56(10):485–8.

[9]   Wang S, Li ZP, Wu JH, Wang Z. Accelerated near-field algorithm of sparse apertures by non-uniform fast Fourier transform. Opt Express. 2019;27(14):19102–18.

[10] Jayanthi S, Vennila C. Advanced satellite image classification of various resolution image using a novel approach of deep neural network classifier. Wirel Personal Commun. 2018;104(1):357–72.

[11] Wu S, Zeng W, Chen H. A sub-pixel image registration algorithm based on SURF and M-estimator sample consensus. Pattern Recognit Lett. 2020;140(8):261–6.

[12] Jia D, Wu S, Zhao M. Dense matching for wide baseline images based on equal proportion of triangulation. Electron Lett. 2019;55(7):380–2.

[13] Paul DS, Divya SV, Pati UC. Structure tensor based SIFT algorithm for SAR image registration. IET Image Process. 2019;14(11):929–38.

[14] Ma W, Wu Y, Liu S, Su Q, Zhong Y. Remote sensing image registration based on phase congruency feature detection and spatial constraint matching. IEEE Access. 2018;6(1):77554–67.

[15] Dixit A, Bag S. Composite attacks-based copy-move image forgery detection using AKAZE and FAST with automatic contrast thresholding. IET Image Process. 2020;14(17):4520–42.

[16] Guo Z, Chen S, Liu H, Yang Q, Yang Z. A fast algorithm for optimal power scheduling of large-scale appliances with temporally-spatially coupled constraints. IEEE Trans Smart Grid. 2020;11(99):1136–46.

[17] Cai JF, Liu H, Wang Y. Fast rank-one alternating minimization algorithm for phase retrieval. J Sci Comput. 2019;79(1):128–47.

[18] Bibin S, Glittas AX, Sellathurai M, Lakshminarayanan G. Reconfigurable 2, 3 and 5-point DFT processing element for SDF FFT architecture using fast cyclic convolution algorithm. Electron Lett. 2020;56(12):592–4.

[19] Wang W, Fu Y, Dong F, Li F. Semantic segmentation of remote sensing ship image via a convolutional neural networks model. IET Image Process. 2019;13(6):1016–22.

[20] Karakuş O, Kuruoğlu EE, Altınkaya MA. Generalized Bayesian model selection for speckle on remote sensing images. IEEE Trans Image Process. 2019;28(4):1748–58.