Adaptive Feature Point Image Registration Algorithm with Added Spatial Constraint Model

Xiao Zhou, Songlin Yu*, Jijun Wang*, Yuhua Chen, Fangyuan Li and Yan Li

Institute of National Defense Engineering, Academy of Military Sciences, Beijing, 100036 China
E-mail: zhouxiao200810@163.com
*Corresponding Author

Abstract

Image data with different spectral features contain different attribute information of a target, which is naturally complementary and can provide more comprehensive and detailed features after registration and fusion. Image registration methods based on point features have the advantages of high speed and precision, and have been widely used in visible light image registration. For registration of multiscale images and those with different spectral characteristics, the precision of these methods is affected by such factors as complex gradient variation. To this end, we add a spatial constraint model to point feature image registration, and improve the method from the aspects of feature point selection, registration, and image conversion parameter calculation. The method is applied to different types of image registration programs, and the results show that it can effectively improve the registration accuracy of multiscale images with different spectral characteristics.

Keywords: Gradient variation, image data, registration.

1 Introduction

Sensors have different imaging mechanisms, and the characteristics of the acquired image data also differ. Visible light images can reveal detailed textures of a target, and their high spatial resolution is conducive to the human eye's recognition and discrimination of scenes and targets [1, 2]; they are therefore widely used in remote sensing. Imaging is more difficult when lighting conditions are poor or a target is obscured by smoke or fog. An infrared image converts the heat radiation of objects in the field of view to greyscale values so as to obtain the target temperature characteristic information. Infrared imaging has unique advantages for thermal targets, as it can penetrate smoke and identify special materials. It is widely used, such as in medical imaging and security monitoring [3]. However, due to hardware and environmental effects, infrared images have high noise, poor visual effects, and low resolution [4]. Due to the different image data characteristics and application scenarios, image data acquired by different types of sensors have natural complementary advantages. An effective fusion of multisource data can improve the ability to detect and analyse target characteristics, and has broad applications such as intelligent monitoring, environmental monitoring, and remote sensing [5–7].

Image registration methods based on point features must extract feature points from the image and register them in order to determine the transformation model between multisource images. Since this type of method is fast, accurate, and performs a small amount of computation, researchers have proposed a variety of point feature registration operators, including the Moravec [8], Forstner [9], SUSAN [10], SIFT [11], and Harris operators [12], and the partial intensity-invariant feature descriptor of Chen et al. [13]. The SIFT operator has scale and rotation invariance, is highly adaptable to affine and grayscale transformations, and is widely used in visible light image registration. When such operators are applied to multimode image registration, they have a low effective registration ratio. This is because the gradient direction of the local area of the image is different from that of the visible light image, and the gradient change and intensity distribution are also inconsistent [14, 15]. To this end, we propose an adaptive image registration method with spatial constraints for automatic and accurate registration of images with different spectral characteristics and on different scales.

We discuss the key technologies of point feature-based registration, and introduce the algorithm flow and the proposed improvement method. The method is experimentally tested for registration of different types of

images. We conclude with an overview of the method and our experimental results.

2 Key Technology of Point Feature Registration

2.1 Point Feature Detection

We determine whether a point chosen for analysis is at the maximum of the local curvature by analysing the greyscale variation of the pixels in various directions. The structure matrix of an image I is

$$N = \sum_{u,v} w(u,v) \begin{bmatrix} I_u I_v & I_u I_v \\ I_u I_v & I_v I_v \end{bmatrix} \tag{1}$$

where $w(u,v)$ is the Gaussian window function; and I_u, I_v are the derivatives of image coordinates (u,v) along the directions of the two respective axes. Let λ_1, λ_2 be the eigenvalues of N, and define it as the cost function, where k is an empirical constant ranging from 0.04 to 0.06. When R exceeds the threshold, the point is considered to be at the maximum of the local curvature change, and it is extracted as a feature point. To improve the detection accuracy, a multiscale corner detection method is used to extract image feature points, and the original image is subjected to fuzzy down sampling using a two-dimensional Gaussian kernel function to construct an image Gaussian pyramid. Adjacent images are subsequently used to construct the Gaussian difference image. By finding the second-order derivative of the Gaussian difference image, feature points at different scales are detected.

Figure 1 shows multiscale and single-scale Harris corner detection results. This picture is a book cover, which contains text information of various features. Using multiscale feature point detection, one can effectively recognize the feature points of most of the text and the background of the book. A total of 1111 feature points are extracted. By detecting the single-scale feature points, some of the more obvious text feature points are recognized, and 200 feature points are extracted. Therefore, in extracting feature points, multiscale extraction results are more beneficial for subsequent feature point registration.

2.2 Feature Descriptor Solution

After a feature point is extracted, its feature descriptor must be solved for feature point registration. A histogram is often used to obtain statistical data

(a) Multiscale   (b) Single-scale

Figure 1   Typical feature point detection results at different scales.

of gradient directions and amplitudes corresponding to the pixels in the neighbourhood centred on the feature point. The histogram is then interpolated and fitted, and the three amplitudes near the maximum are used to calculate the maximum gradient value by parabolic interpolation. This is used as the direction of the feature point. The gradient direction of the main direction of the feature points corresponding to a 4×4 window of the feature point neighbourhood is calculated. The histogram statistical method is similarly used to calculate the gradient direction and amplitude of each region. Each gradient histogram is divided into eight directional intervals between 0 and $360°$, i.e., at intervals of $45°$. Each neighbourhood window contains the gradient intensity information of eight directions, and the descriptor of the feature point is a 128-dimensional vector. The registration of the feature points is based on the similarity of the feature point descriptors.

3 Improved Feature Point Registration Method

When a commonly used registration method based on point features is applied to multimodal image registration, the effective registration ratio is often low. We propose an adaptive image registration method with an added spatial constraint model, which can be used for automatic precision registration of images on different scales and with different spectral characteristics. Feature points are extracted using the improved SIFT algorithm, and the feature point descriptor is constructed by combining the characteristics of the multimodal image. A spatial constraint model is added to improve registration

accuracy, and the registered feature point information is used to register multimodal images using the image mapping relationship under the Gaussian statistical model.

3.1 Improved Feature Descriptor

Due to the different light and shade exhibited by multimodal images, the gradient directions of corresponding regions also differ. The changes in feature descriptors are relatively more complex compared with those of visible light images. The changes in the direction of feature descriptors between visible light images are relatively small, whereas images with different spectral features cannot easily meet the requirements between visible light images.

Considering the possible inconsistency of image gradient changes and intensity distributions between multimodal images, we set up two feature descriptors, one with eight direction intervals in the range 0–2π, and the other combining the reversed gradient as one direction, i.e., four direction intervals are set in the range 0–π. The similarities of the two descriptors are combined when feature points are registered. This can reduce the registration error caused by gradient flip.

3.2 Registration Method with Geometric Constraints

When collecting data on a target, a spatial geometric relationship exists between sensors, and therefore between the targets in a multisource image, i.e., the direction and distance of the feature points to be registered meet a certain spatial constraint. For this reason, we introduce a spatial constraint model to feature point registration. We statistically analyse the horizontal, vertical, and feature distances between the feature points registered with the aid of feature descriptors. We determine the respective expected values $\bar{X}$, $\bar{Y}$, $\bar{R}$ of the absolute difference, calculate the corresponding variance δx, δy, δr between the feature points, and analyse, point by point, to determine whether the feature point pair satisfies the geometric constraints

$$|\Delta X - \bar{X}| \leq 2\delta x$$
$$|\Delta Y - \bar{Y}| \leq 2\delta y$$
$$|\Delta R - \bar{R}| \leq 2\delta r \tag{2}$$

where ΔX, ΔY, and ΔR are the respective horizontal, vertical, and distance differences.

If a candidate feature point pair satisfies the above spatial constraints, it is kept for subsequent calculation of image transformation model parameters, otherwise it is deleted. This more stringent spatial constraint will keep the optimally registered feature point pairs, but to ensure that there are enough feature points in calculations for image registration parameter optimization, the geometric constraint is relaxed when there are fewer than five selected feature points after geometric constraints. For example, twice the variance may be changed to three times, and re-registration is implemented with the relaxed geometric constraints until sufficient feature points are selected.

3.3 Registration Method with Geometric Constraints

Image registration solves the transformation model G between pixel values $I_1(x, y)$ and $I_2(x, y)$ at image (x, y) of multisource images, i.e.,

$$I_1(x, y) = G(I_2(x, y)). \tag{3}$$

Considering the existence of rotation, translation, and scaling between images of different scales, an affine transformation is used as the transformation model between image registrations,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} + T \tag{4}$$

where A is the affine transformation matrix, including the scaling and rotation parameters of the transformation model, and T is the translation matrix, representing the translation parameters.

Due to differences in image resolution and sensor field of view, a certain disturbance difference exists between the registered feature points. This should conform to a Gaussian distribution with zero expectation, i.e.,

$$I_1(x, y) - G(I_2(x, y)) \sim \zeta(0, \sigma^2 I) \tag{5}$$

where $\sigma^2 I$ is the covariance and I is the identity matrix. Assuming that the feature point dataset used for registration is independent and identically distributed, the likelihood function between the two is

$$\mathrm{p}(Y|X, \theta) = \prod_{i=1}^{I} \mathrm{p}(Y_i|X_i, \theta). \tag{6}$$

Since the feature points of dataset θ conform to a Gaussian distribution, the transformation parameters are iteratively calculated to maximize the
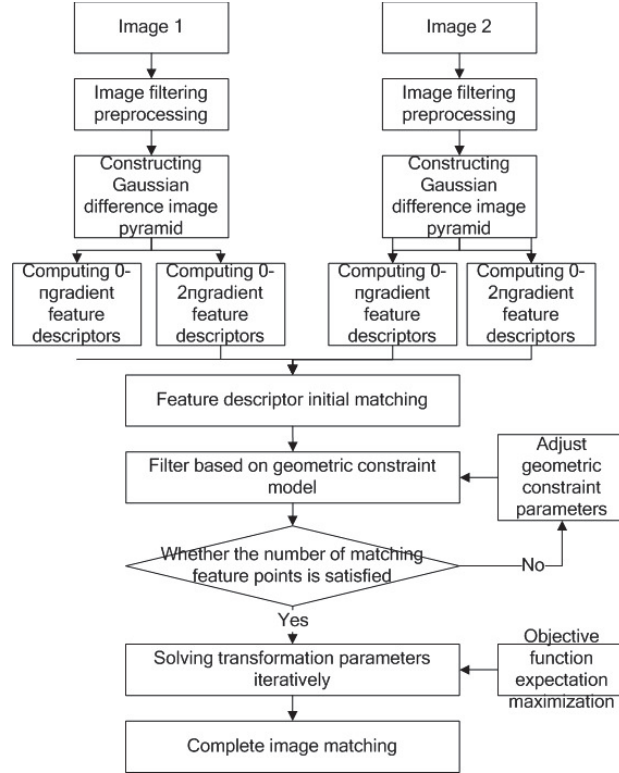
Figure 2   Data processing flowchart of proposed algorithm.

likelihood function between the data,

$$\theta^{new} = \arg \max_{\theta} \mathrm{p}(\theta, \theta^{old}). \tag{7}$$

Finally, the transformation parameters can be used to complete the registration between the pixels in the image. Figure 2 shows a flowchart of the algorithm.

4 Analysis of Experimental Results

To verify the feasibility of the proposed algorithm, we performed registration experiments on visible light images, and visible and near-infrared light images of different target scenes. Statistical analyses were performed on the results.

4.1 Visible Light Image

The visible light image registration test was performed on book cover photos taken at different angles, containing Chinese, English, and other feature targets. Feature point registration was implemented with and without geometric constraints, with the results shown in Figure 3. A total of 375 feature points were registered without geometric constraints, including a number that were mis-registered. The characteristics of the connections between mis-registered points are different from those between correctly registered feature points; they cross with the connections between correctly registered points. These errors occur due to the complexity of the character shape, and because similar feature descriptors can easily appear and cause registration error. With geometric constraints, there were 358 registered feature points, as shown in Figure 4. Most were correctly registered, but incorrectly registered points still appeared near the top of the letter "J" because there were two very similar local regional image features in this letter, and the differences
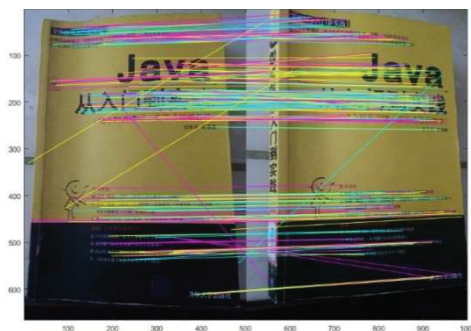

Figure 3   Feature point registration results without geometric constraints.

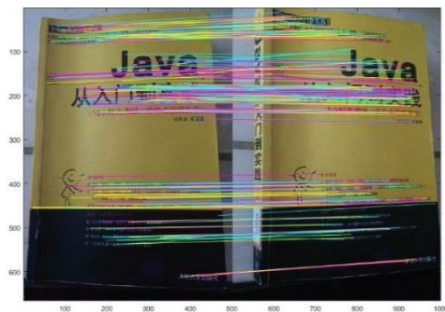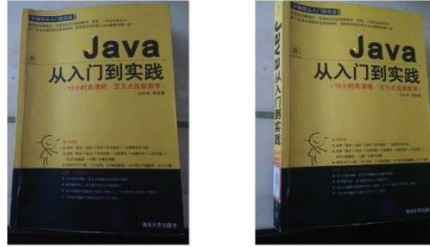**Figure 4**   Registration results after geometrically constrained model.

(a) Selected feature points   (b) Corresponding feature points after solution

Figure 5   Feature point registration results.

between feature descriptors and between spatial geometric relationships were relatively small.

Since the principle of maximum likelihood is adopted in the calculation of image transformation parameters, and the method takes into account the statistical information of global feature points, the influence of mis-registered feature points can be effectively reduced during feature parameter calculation. To check the image results after registration, we chose five feature points in the letter "J" of a registered image, and compared them with the position information of five corresponding feature points in another image, as shown in Figure 5. The comparison showed that the mis-registered feature points were effectively corrected, and that all five pixels in the letter "J" were correctly registered.

4.2 Visible Light and Near-infrared

The effectiveness of the proposed algorithm was validated by testing two sets of visible and near-infrared light images of different scenes. One group was composed of targets with similar characteristics to those of trees and vegetation, and the other set had greater differences, including vehicles, buildings, and calibration boards.

The feature point registration results of the first group of images are shown in Figures 6 and 7, from which it can be seen that near-infrared images have smaller fields of view than visible light images. Figure 6 shows registration results without geometric constraints for 56 feature points. The images contain many mismatched results; for example, the tree branches at the upper-left of the visible light image are registered with the bottom area of the tree trunk in the near-infrared image, and the turf in the visible light is registered
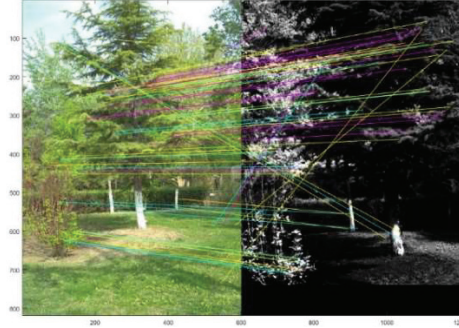
Figure 6   Registration results without geometric constraints.
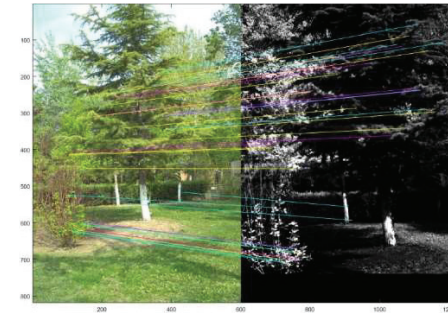

Figure 7   Registration results with geometric constraints.

with tree branches in the near-infrared image. Figure 8 shows the registration results with geometric constraints, where 36 feature points are registered. The proposed method filtered out most of the mis-registered points, but there are still mis-registrations. For example, the trunks of the different trees are registered because the images have similar features. This is similar to the mis-registration of "J" in visible light, and must be deleted through the global feature of the feature point when the registration parameters are calculated. Similarly, five registered image feature points were selected and the registration results analysed, as shown in Figure 8. The results show that the correspondence relationship between the mis-registered tree feature points have been effectively corrected, and the pixels at the positions of the five feature points are accurately registered.

The registration results of the second group of images are shown in Figures 9 and 10, where there are 35 registration feature points without geometric constraints. The feature points of the calibration checkerboard and

(a) Selected feature points   (b) Corresponding feature points after solution

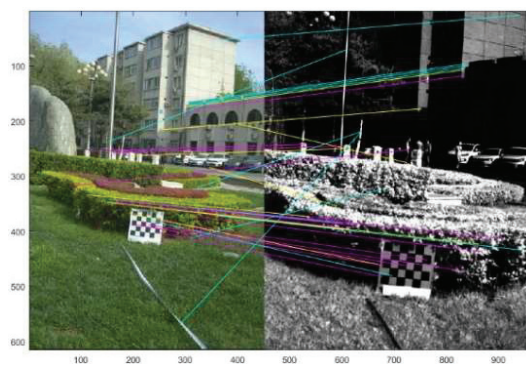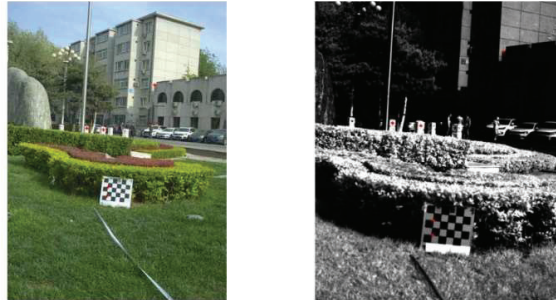**Figure 8**   Comparison results of feature points after registration.

Figure 9   Registration results without geometric constraints.

Figure 10   Registration results with geometric constraints.

(a) Selected feature points   (b) Corresponding feature points after solution

Figure 11   Comparison results of feature points after registration.

the boundary area of the building are all correctly registered, but there are still many mis-registration results, e.g., the water hose on the lawn in the visible light image is registered with the tree trunks in the near-infrared image. There are 16 registered feature points with geometric constraints, all accurately registered, and there are no mis-registered points. Again, five registered image feature points were chosen to examine the final registration results, including two calibration board feature points, two landmark feature centre points, and one building corner point, as shown in Figure 11, which shows that they are all accurately registered. According to the registered feature points on the calibration board, the image accuracy after registration is better than two pixels.

5 Conclusions

In the registration of images on different scales and with different spectral characteristics, the multiscale feature point detection method can effectively increase the number of single-scale feature points to be extracted. However, due to different light and dark presentation between images, there are gradient changes and intensity distribution inconsistencies. This leads to a low effective registration ratio for the point feature registration method. Feature point registration efficiency can be effectively improved by constructing multimode descriptors.

Using the spatial geometric relationship between the targets in an image, a spatial constraint model can be added to effectively improve feature point registration accuracy. Since the changes in spatial geometric relationships are

small for some of the feature points, they cannot be effectively discriminated through statistical models alone. In solving registration parameters between images, one can exploit the fact that the feature point dataset after registration follows a Gaussian distribution, and thereby make use of the maximum likelihood function to reduce the effect of some mis-registered feature points, so as to achieve high-precision image registration.

The proposed algorithm can not only improve the registration accuracy between visible images, but also between visible and infrared images. From the analysis of the registration results of different types of images, it can be seen that the proposed method can automatically register images with different spectral characteristics acquired by multiple types of sensors at different angles, and the results can reach the pixel level.

References

[1] M. Ding et al., "Infrared target detection and recognition method in airborne photoelectric system," *Journal of Aerospace Information Systems*, vol. 16, no. 3, pp. 1–13, Feb. 2019.

[2] C. Romain et al. "Type-II superlattice-based extended short-wavelength infrared focal plane array with an AlAsSb/GaSb superlattice etch-stop layer to allow near-visible light detection," *Optics Letters*, vol. 42, no. 21, pp. 4299–4302, Oct. 2017, doi: 10.3390/s150817944.

[3] J. Han, E. J. Pauwels, and P. D. Zeeuw, "Fast saliency-aware multi-modality image fusion." *Neurocomputing*, vol. 111, pp. 70–80, July, 2013, doi: 10.1016/j.neucom.2012.12.015.

[4] C. S. Martín, "Fusion of visible and thermal descriptors using genetic algorithms for face recognition systems," *Sensors*, vol. 15, no. 8, p. 17944, doi: 10.3390/s150817944.

[5] A. Dehzangi et al., "nBn extended short-wavelength infrared focal plane array," *Optics Letters*, vol. 43, no. 3, pp. 591–594, Feb. 2018.

[6] J. Xin et al., "A survey of infrared and visual image fusion methods," *Infrared Physics & Technology*, vol. 85, pp. 478–501, doi: 10.1016/j.infrared.2017.07.010.

[7] M. Uzair et al., "Periocular region-based person identification in the visible, infrared and hyperspectral imagery," *Neurocomputing*, vol. 149, no. 1, pp. 854–867, Jan. 2015.

[8] H. P. Moravec, "Towards Automatic Visual Obstacle Avoidance," *Proc. Int. Joint Conf. on Artificial Intelligence, 1977,* pp. 584–585.

[9] H. E. Yinan et al., "Integration of Forstner and NCC constraint for UAV image registration," *Remote Sensing for Land & Resources*, vol. 27, no. 1, pp. 48–54. doi: 10.6046/gtzyyg.2015.01.08.

[10] H. Bay et al., "Speeded-up robust features (SURF)," *Computer Vision & Image Understanding*, vol. 110, no. 3, pp. 346–359, Jun. 2008.

[11] C. Harris and M. Stephens "A combined corner and edge detector," *Proceedings of Proc. Alvey Vis. Conf.*, to be published, doi: 10.5244/ C.2.23.

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints." *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–114, Nov. 2004.

[13] Q. S. Chen and M. Defrise, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 16, no. 12, pp. 1156–1168, Dec. 1994.

[14] J. Ma et al, "Robust Point Matching via Vector Field Consensus," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1706–1721, Feb. 2014.

[15] W. Han and W. Ming, "CCD-IR image registration based on adaptive feature point detection," *Journal of Image and Graphics*, 2017.

Biographies

Xiao Zhou graduated from the School of Surveying and Mapping of Wuhan University in 2015, with Ph.D. in engineering. He is now an engineer for the National Defense Engineering Research Institute of the Academy of Military Sciences. His research interests include environmental perception, artificial intelligence, and digital image processing.

Songlin Yu graduated from Nanjing University of Technology with a Master's degree. He is now an engineer for the National Defense Engineering Research Institute of the Academy of Military Sciences. His research interests include environmental perception, and digital image processing.
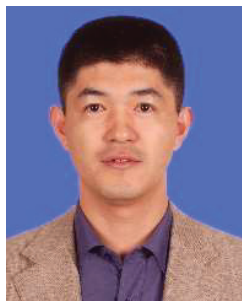
Jijun Wang graduated from Nanjing University of Technology with a Master's degree. He is now a senior engineer for the National Defense Engineering Research Institute of the Academy of Military Sciences. His research interests include environment perception, big data processing, and artificial intelligence.

Yuhua Chen graduated from Nanjing University of Technology with a Ph.D. degree. He is now a senior engineer for the National Defense Engineering Research Institute of the Academy of Military Sciences. His research interests include precision optical image processing and artificial intelligence.

Fangyuan Li graduated from Beijing University of architecture and Engineering with a Master's degree. She is now an engineer of the National Defense Engineering Research Institute of the Academy of Military Sciences. Her research interests include environmental perception and digital image processing.

Yan Li graduated from Central China Normal University with a Master's degree. She is now an engineer for the National Defense Engineering Research Institute of the Academy of Military Sciences. Her research interests include surveying and mapping engineering, image processing, and big data analysis.