REGISTRATION OF MULTITEMPORAL GF-1 REMOTE SENSING IMAGES WITH WEIGHTING PERSPECTIVE TRANSFORMATION MODEL

Ruitao Feng[1], Xinghua Li[2], Wenli Zou[3], Huanfeng Shen[1]

[1] School of Resource and Environmental Sciences, Wuhan University, P. R. China
[2] School of Remote Sensing and Information Engineering, Wuhan University, P. R. China
[3] Faculty of Information Engineering, China University of Geosciences (Wuhan), P. R. China

ABSTRACT

Registration is a classical problem in the application of remote sensing images. The existing methods prefer to fit the relationship between target and source images with the same model on the whole. In fact, the geometrical relationship between two images is not always consistent, especially for the wide-field-viewed images of GaoFen-1 (GF-1) launched by the China Aerospace Science and Technology Corporation (CASC) in April 2013. Generally, The existing methods didn't take the local deformation into consideration. Towards this end, we solve the problem with three stages in this paper. Firstly, the coarse registration obtains the integral perspective transformation model of images. Secondly, the fine registration partitions image into many blocks and improves the relationship of every block with the inverse distance weighting (IDW) function. Finally, the coordinate transformation and resampling are the final step. Compared to other methods, the experiments demonstrate that the proposed algorithm is capable of generating satisfied results which are robust against deformation at local area.

Index Terms— Feature point, feature line, GF-1 images, IDW, perspective transformation model

1. INTRODUCTION

Aligning different images plays an important role in image processing, such as mosaic [1], image fusion [2], multi-temporal images analysis [3] and so on. The current methods could be divided into two categories: intensity-based and feature-based.

The intensity-based approaches usually calculate the similarity of two images with an optimized algorithm pixel by pixel. Generally, the mutual information (MI) [4], cross-correlation (CC) [5] and sequential similarity detection algorithm (SSDA) [6] are often used as criteria. The higher the similarity, the better the registration result. However, this class of methods requires higher quality images and is easily influenced by external conditions. Comparing to digital images and photographs, there are abundant objects in remote sensing images and the field of view is wider. The computation is so heavy that the intensity-based methods are not widely used in remote sensing images. The feature-based methods describe images with extracted features which are mostly usable for remote sensing fields. Point, polyline and polygon are common features for registration. Feature line and polygon are less-using because the matching is still a challenging problem. Points are the mostly used feature. The extraction of feature point has a large number of methods, such as SIFT [7], SURF [8], KAZE [9] and so forth. SIFT is a widely used operator which extracts more feature points. SURF is based on SIFT, which improves the efficiency at the cost of losing accuracy. KAZE extracts more feature points than SURF by constructing nonlinear space. However, it is difficult to acquire feature points in low texture region [10]. At the moment, dual feature connecting point and polyline together is the first choice to estimate the transformation model [11]. In contrast of intensity-based methods, feature-based methods are more convenient and less time-consuming which is suitable for real-time processing of remote sensing images.

Since the accuracy of extracting features is high enough, the key point of registration focuses on the estimation of the geometrical relationship. Affine transformation model is widely used which is the representative of linear models. Nevertheless, it is not appropriate to correct the deformation of shapes. The perspective transformation model which is the representative of nonlinear transformation model is put forward to address the problem. However, a perspective model on the whole image lacks consideration of local deformations. Thus, dividing images and estimating the relationship locally are better. In the seminal work of [12], the author partitions source image into a grid of $C_1 * C_2$ cells and computes the transformation model of every cell with the support of a weighting function which is similar to Gaussian kernel function. However, it mainly aims to align digital images, which ignores the geographical law of objects.

In this paper, we made two main contributions. On the one hand, the perspective transformation model is enforced on the registration of GF-1 images, which utilizes the feature points and lines. On the other hand, in consideration of the inconsistence of local deformations, the transformation model is adjusted by weighting blocks using the inverse distance weighting (IDW).

The rest of the paper is organized as follows. In Section 2, we give a description of registration from three stages, in which the attention is paid on the perspective transformation model, IDW and the solution for parameters. After that, we verify the proposed method by real experiments in Section 3. At last, a conclusion for the current work is provided in Section 4.

2. METHODOLOGY

The typical image registration generally includes feature extraction, feature matching, transformation model construction, coordinate transformation and resampling. Based on the fundamental steps, we propose a systematic processing method showed in Fig. 1.

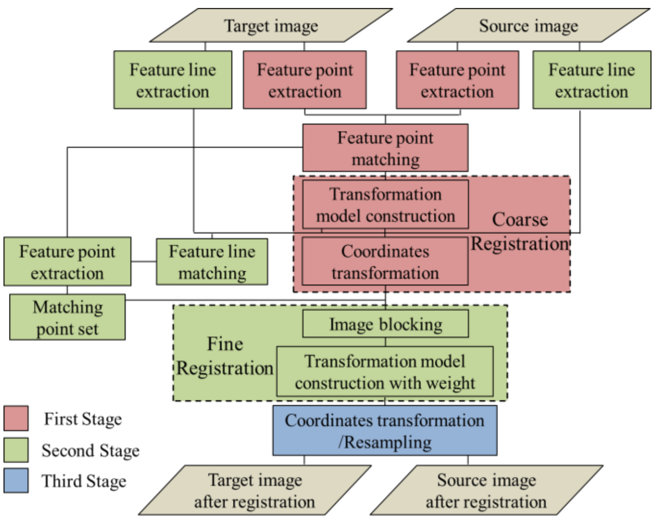


Fig. 1 Flow chart of proposed method

2.1. Coarse registration

Feature points are detected and described by SIFT which is the most robust operator. Comparing the distance of the closest neighbor to that of the second-closest neighbor is an effective measure to obtain matching pairs [7]. Furthermore, on account of clouds and highlight buildings, there are many wrong-matching points. It is necessary to remove them and reserve inner points with random sample consensus (RANSAC) [13]. Then we estimate the perspective transformation model with the support of inner points for whole image.

A perspective transformation model H aims to map coordinates in target image to correspondence in source image.

$$X' = HX \tag{1}$$

For ease of calculation, let $X = (x, y, 1)$ and $X' = (x', y', 1)$ be $X = (x, y)$ and $X' = (x', y')$ in homogeneous coordinates.

Direct linear transformation (DLT) is a basic method to estimate H from a set of feature points [14]. If the transformed point in target image is coincide with original corresponding point in source image, Eq.(1) is rewritten as $X' \times HX = \vec{0}$ and linearized

$$\begin{bmatrix} 0 & -X^T & y'X^T \\ X^T & 0 & -x'X^T \\ -y'X^T & x'X^T & 0 \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \end{bmatrix} = \vec{0} \tag{2}$$

where $H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}^T$, and h_i is the row of H, $i = 1, 2, 3$.

Actually, DLT estimates H as

$$\tilde{h} = \arg\min_h \sum_{i=1}^{N} \|m_i h\|^2 = \arg\min_h \|Mh\|^2, \ s.t.\ \|h\| = 1 \tag{3}$$

where N is the number of points. After transforming four vertex coordinates of source image, we could get the result of the coarse registration.

2.2. Fine registration

Aiming to enlarge the feature points in low texture region, center points of matched feature lines are added to the matching point set.

Since the relationship of source and target image varies from one region to another, we divide the result of first stage into blocks and develop a block-dependent transformation model. Not only points in the block but also points distributed around it could affect the accuracy of calculation. Thus, weighting function is a good choice to solve the problem whether the point is contributed or not.

2.2.1. Inverse distance weighting (IDW)

IDW is one of the most frequently used deterministic models in spatial interpolation [15]. It is developed from the First Law of Geography, specially about that everything is related to everything else, but near things are more related than distant things [16]. Therefore, we think that the transformation model is not only determined by feature points in blocks, but also around the blocks. The degree is established by IDW and the weight is calculated as follows:

$$w_i = \frac{1 \big/ \sqrt{(\mathrm{x}_i - \mathrm{x}_*)^2 + (\mathrm{y}_i - \mathrm{y}_*)^2}}{\sum_{i=1}^{N} 1 \big/ \sqrt{(\mathrm{x}_i - \mathrm{x}_*)^2 + (\mathrm{y}_i - \mathrm{y}_*)^2}} \tag{4}$$

where (x_i, y_i) is the ith coordinate of feature points in target image. And (x_*, y_*) is the center point of block.

2.2.2. The solution of weighting perspective transformation model

For each block, the distribution of weight is arranged as in Fig. 2.

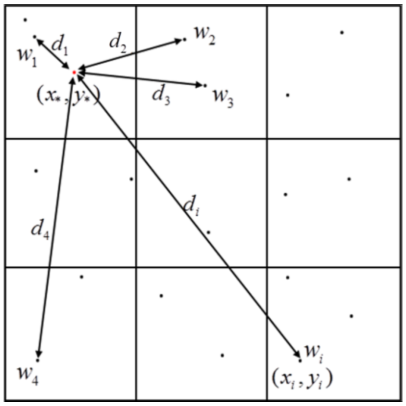


Fig. 2 the distribution of weight

And the equation to estimate the transformation model is written as

$$\tilde{h} = \arg\min_h \sum_{i=1}^{N} \|w_i m_i h\|^2 = \arg\min_h \|WMh\|^2 \qquad (5)$$

where w_i is the weight of the ith point calculated by Eq.(4). The equation could be established for each pair of points. By stacking $w_i m_i$ for all points, the solution is simply the least significant right singular vector of WM .

2.3. Coordinate transformation and resampling

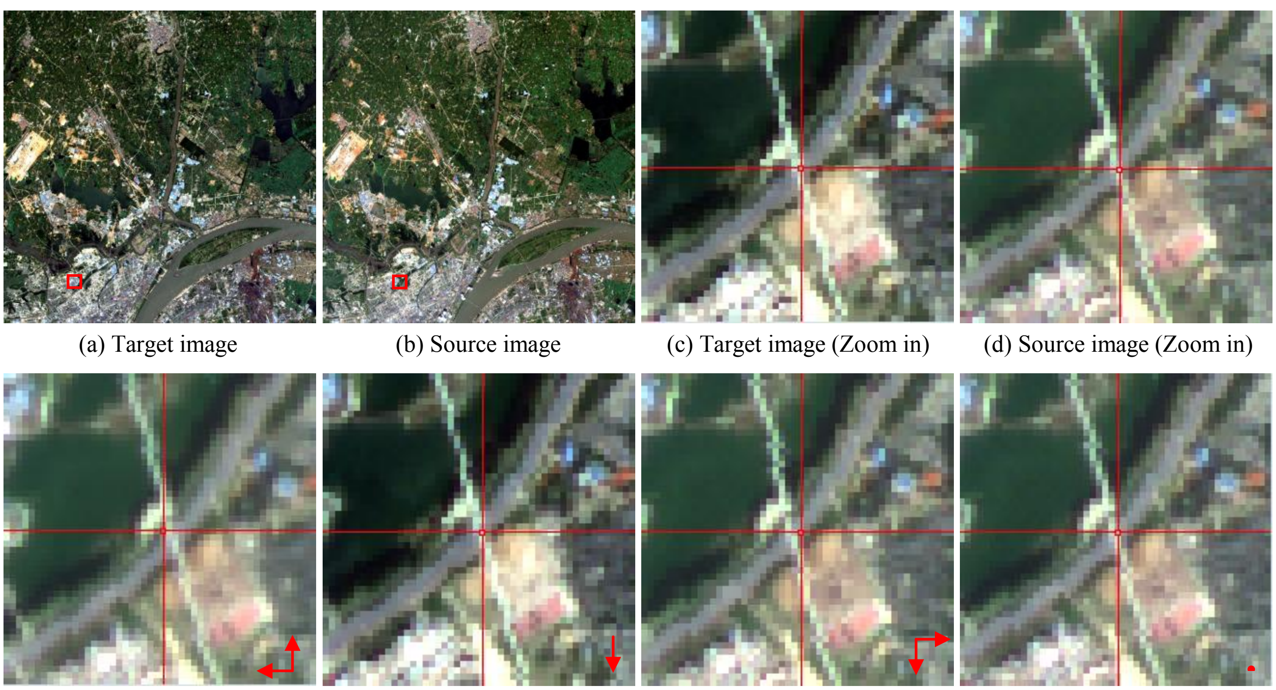
Coordinates in source image is transformed, and resampling is used to calculate the intensity of point located in non-integer coordinates.

3. EXPERIMENTS

In this paper, experimental images are the level-1A data with the process of homogenized radiation calibration from the wide field of view (WFV) imaging system onboard the Chinese GF-1 optical satellite [17]. The resolution is 16-m and the width is 800-km. Two multi-temporal sets of images are chosen for verifying the proposed method. The size of images is 2000*2000 pixels. All of the images are divided into 5*5 blocks (400*400). The weight of feature points far from the center of block is insignificant. They may affect the stability of Eq. (3). Thus, we offset the weight with an empirical value $\sigma = 0.004$.

Experiment 1: The target image is taken on August 3, 2015 and the source image is taken on August 7, 2015, showed in **Fig. 3**. It is Wuhan, China.

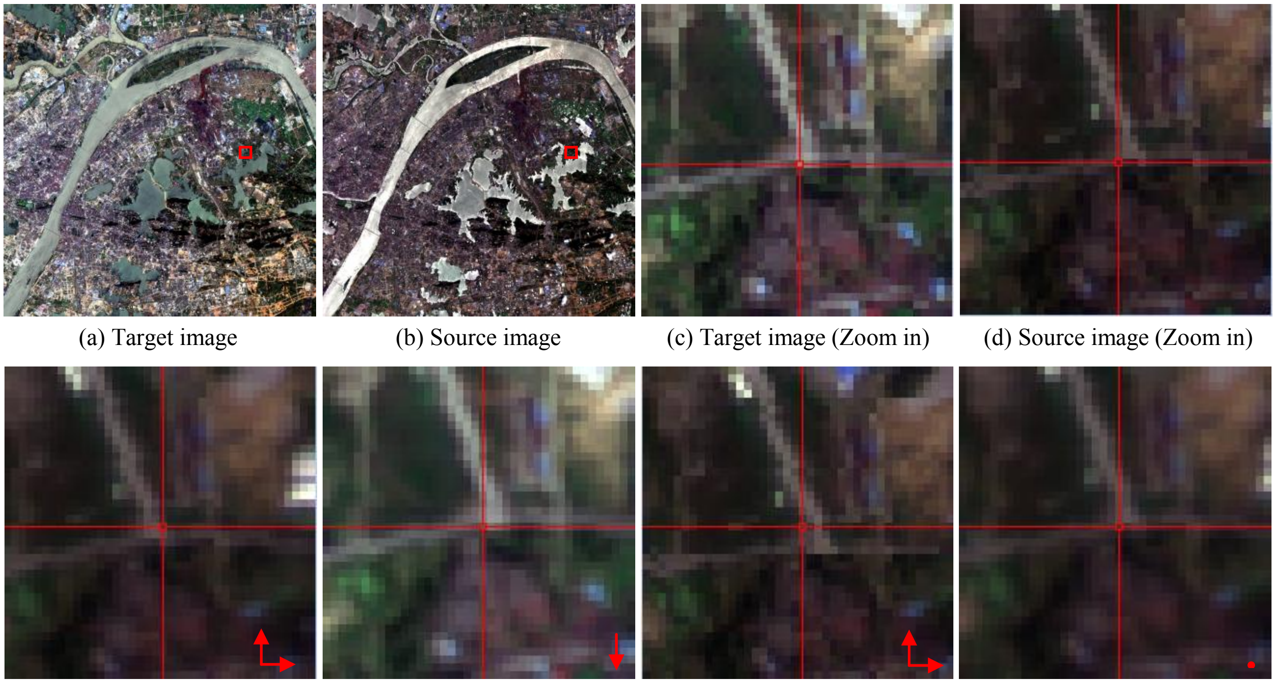
Experiment 2: The target image is taken on July 29, 2016 and the source image is taken on June 14, 2016, showed in Fig. 4. It is Wuhan, China.



(a) Target image (b) Source image (c) Target image (Zoom in) (d) Source image (Zoom in)

(e) ENVI (f) Total-Perspective (g) APAP (h) Ours

Fig. 3 Comparisons of registration for Experiment 1 (the red arrows is the location of results relative to the target image).



(a) Target image (b) Source image (c) Target image (Zoom in) (d) Source image (Zoom in)

(e) ENVI (f) Total-Perspective (g) APAP (h) Ours

Fig. 4 Comparisons of registration for Experiment 2 (the red arrows is the location of results relative to the target image).

For both Fig. 3 and Fig. 4, (a) and (b) are the original images. In order to observe the result, the red block in original images and results are zoomed in showed in (c) and (d). (e) is the result of ENVI. (f) is the red block aligned with a perspective transformation model on whole image. (g) is the result of APAP [12]. The image is partitioned into 100*100 cells in the proposed range. (h) is the result of our method. Our result is better than the results of ENVI, total-perspective and APAP. In order to observe the results conveniently, red arrows mean that there are translations in horizontal and vertical directions. Red points are labeled on the images to represent no translations.

The root-mean-square error (RMSE) is used as quantitative assessment criteria,

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left\|H(X_i) - X_i^{'}\right\|^2} \tag{6}$$

where M is number of points, H is the transformation model of corresponding point, and X *and* X' are the coordinates in target and source image. With no ground control points, 20 uniformly distributed points are extracted to evaluate the result of registration.

Table 1 RMSE of registration results

	Fig. 3	Fig. 4
ENVI	3.9609	10.8036
Total-Perspective	1.8165	1.8575
APAP	0.6215	0.8242
Ours	**0.6118**	**0.7736**

The smaller the RMSE, the better the registration. The result of ENVI which estimates the relationship between two images with affine model is worse than others. Our method outperforms other methods, in which RMSE is less than one pixel. In this way, our method could offer precise results for the further research.

4. CONCLUSION

To satisfy high-precision alignment of GF-1 images, this paper proposes a three-stage registration method using joint feature points and feature lines. The coarse registration is the first stage in which the same perspective model is built on the two images. Then the fine registration improves the transformation model block by block with the support of IDW. Thirdly, the coordinate transformation and resampling are applied to fulfill the registration. Comparing to the other three methods, our method takes the local deformation into consideration and is more accurate in both visual and quantitative results.

5. ACKNOWLEDGMENTS

This work was supported by the Fundamental Research Funds for the Central Universities of China under Grant No. 2042017kf0034, and the Open Research Fund of the Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences under Grant No. 2016LDE004. The authors would also like to thank the anonymous reviewers.

6. REFERENCES

[1] X. Li, N. Hui, H. Shen, Y. Fu, and L. Zhang, “A robust mosaicking procedure for high spatial resolution remote sensing images,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 109, pp. 108–125, 2015.

[2] H. Shen, X. Meng, and L. Zhang, “An Integrated Framework for the Spatio–Temporal–Spectral Fusion of Remote Sensing Images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7135–7148, 2016.

[3] C. Zeng, H. Shen, and L. Zhang, “Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method,” *Remote Sensing of Environment*, vol. 131, pp. 182–194, 2013.

[4] A. Colerhodes, I. Zavorin, and J. L. Moigne, “Mutual information as a similarity measure for remote sensing image registration,” in *Proceedings of SPIE-The International Society for Optical Engineering*, 2001, pp. 51–61.

[5] A. Martinez, J. Garcia-Consuegra, and F. Abad, “A correlation-symbolic approach to automatic remotely sensed image rectification,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 1999, pp. 336–338.

[6] D. I. Barnea and H. F. Silverman, “A Class of Algorithms for Fast Digital Image Registration,” *IEEE Transactions on Computers*, vol. 100, no.2, pp. 179–186, 1972.

[7] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[8] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (SURF),” *Computer Vision and Image Understanding*, vol. 110, no.3, pp. 346–359, 2008.

[9] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, “AZE Features,” in *12th European Conference on Computer Vision (ECCV)*, pp. 214–227.

[10] A. Goshtasby, G. C. Stockman, and C. V. Page, “A Region-Based Approach to Digital Image Registration with Subpixel Accuracy,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, no.3, pp. 390–399, 1986.

[11] S. Li, L. Yuan, J. Sun, and L. Quan, “Dual-Feature Warping-Based Motion Model Estimation,” in *International Conference on Computer Vision (ICCV)*, 2015, pp. 4283–4291.

[12] J. Zaragoza, T. J. Chin, M. S. Brown, and D. Suter, “As-Projective-As-Possible Image Stitching with Moving DLT,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1285–1298, 2014.

[13] J. Gao, S. J. Kim, and M. S. Brown, “Constructing image panoramas using dual-homography warping,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2011, pp. 49–56.

[14] Hartley R, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge university press, 2003.

[15] G. Y. Lu and D. W. Wong, “An adaptive inverse-distance weighting spatial interpolation technique,” *Computers & Geosciences*, vol. 34, no. 9, pp. 1044–1055, 2008.

[16] W. R. Tobler, “A Computer Movie Simulating Urban Growth in the Detroit Region,” *Economic Geography*, vol. 46, no. sup1, pp. 234–240, 1970.

[17] Li, Z., Shen, H., Li, H., Xia, G., Gamba, P., & Zhang, L. (2017). Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery. Remote Sensing of Environment, 191, 342-358.