Study on Application of Scale Invariant Feature Transform Algorithm on Automated Geometric Correction of Remote Sensing Images

Hui Deng[1,2], Limin Wang[1,2], Jia Liu[1,2], Dandan Li[1,2], Zhongxin Chen[1,2], Qingbo Zhou[1,2]

[1] Key Lab. of Resources Remote Sensing & Digital Agriculture, Ministry of Agriculture, 10081, Beijing, P. R. China

[2] Institute of Agriculture Resources and Regional Planning, Chinese Academy of Agricultural Sciences, 10081, Beijing, P. R. China

Abstract. In recent years, data collected from remote sensing satellite and aerophotography have been showing a geometric sequence increase. A method of Scale Invariant Feature Transform (SIFT) algorithm could be employed for the automatic geometric fine correction. This method could avoid the impact of the rotation and zooming of template matching during the image matching process, and it can also save the labor during the image processing operation. Based on the SIFT algorithm, this paper proposes a two-step method, which firstly conducts coarse match on feature points, and then further conducts fine correction on the coarsely matched feature points by using the least squares technique. The result indicates that, this method is an effective automatic matching method for remote sensing images.

Keywords: remote sensing image, automatic geometric correction, SIFT algorithm, least squares

1   Introduction

Along with the rapid increase of huge amount of remote sensing data, the workload of geometric fine correction on manual image matching for the image data collected from different sources and in different time period is huge. Image matching is the preliminary work of a lot of remote sensing applications (such as multi-source data fusion, variation detection, image mosaic, motion detection, and object identification, etc, and its result directly affects the quality of the follow-up works[1]. Currently, the matching technology for the satellite remote sensing images from the same sensor, at different time period, with small angular variation has been relatively mature, and some classical feature extraction algorithms, such as arithmetic operators[2] ,

arithmetic operators[3], are extracted. They proposes a multiscale registration technique using robust Scale Invariant Feature Transform (SIFT) features in Steerable-Domain[4]. The matching the images from multi-sensor with different resolutions is the focus and the difficulty of current researches[5]. Scholars both at home and abroad have conducted extensive in-depth research on this topic, but no perfect method, which could correctly match the multi-source image under different radiation distribution, dimension, rotation, and translation, has been established.

In recent years, scale invariant Feature Transform (SIFT) feature has achieved great success in image matching. The matching method based on the SIFT descriptor has been successfully applied in many fields, such as object identification [6] , motion panorama[7], and Structure From Motion (SFM)[8]. SIFT algorithm was introduced [6] and [9] in 1999 and was well established in 2004. They also conducted experiments on the many representative descriptors under six circumstances, such as different scenes, change in illumination, and geometry deformation of images, resolution difference, rotation, blurring, and image compression. The results indicate that SIFT descriptor has the best performance [10].

2. Introduction to the algorithm

The extraction and matching algorithm based on SIFT point features mainly includes three steps: extraction of point features, computation of feature descriptors, and feature matching from [6] and [11].

The Gaussian smoothing is applied on original image with different standard deviation σ, to obtain the Gaussian differential image. The maximum or minimum gray-scale values on differential images are taken as the feature points. Formula (1) is gaussian smoothing formula, and the extraction process of feature points is illustrated in the Figure 1. In the Figure, each pixel on the Gaussian differential image (expressed by crosshair) is compared with 26 adjacent pixels (expressed by dot) within the 3X3 area of current and upper and lower adjacent dimensions. If the average values of Difference of Gaussian of a pixel is larger or smaller than the adjacent pixels, the pixel is taken as the local extreme point. Then, some local extreme points are eliminated according to the preset threshold value of Difference of Gaussian and edge effect standard, and remaining extreme points are taken as the final SIFT feature points.

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

By taking feature point as the center, the image area with specified height and width is taken, to calculate the gradient direc0 tion and gradient strength of each pixel within the area. The total gradient strength of different gradient directions are calculated and are taken as the eigen vector. The calculation process of the eigen vector is illustrated in the Figure 3. In the Figure, the arrow in the left stands for the gradient strength and direction. The total gradient strengths of 8 directions within the 2×2 sub-areas are calculated separately, so as to obtain an eigen vector with 32 dimensions. The calculation of Lowe's study (Lowe，2004) was on 8 directions within a 4×4 sub-area, so the feature descriptor of each SIFT point feature is the eigen vector of 4×4×8=128 dimensions.
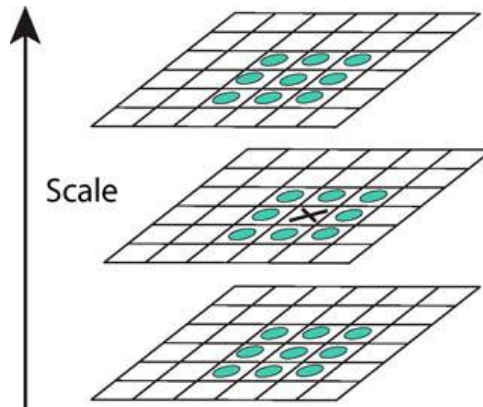
Fig. 1. SIFT point feature extraction process （Lowe，2004）

Euclidean distance of the eigen vector of different feature points of image to be matched and Reference image is calculated, and the feature point with minimum distance is taken as the initial matching point. The error matching points are eliminated according to the nearest neighbor and second nearest neighbor Euclidean distance ratio.

Image gradients   Keypoint descriptor
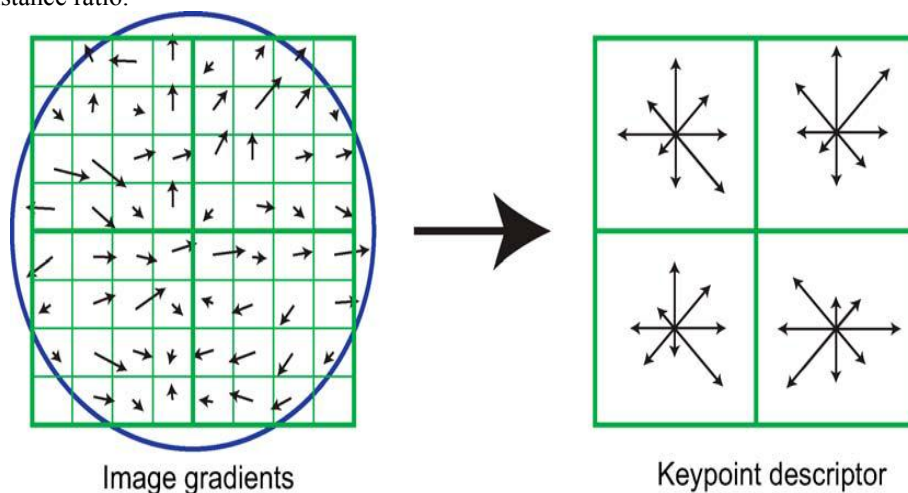
Figure 2. Feature descriptor calculation process (Lowe，2004)

3. Analysis and Conclusion

Firstly, feature point and eigen vector information of images to be corrected and Reference images is calculated. The partitioning technology is used to locate feature point pairs, so as to improve the feature point search efficiency, and then the coarse matching is conducted based on both images to find out the matched feature point pairs on both   images. After the coarse matching, the fine matching on feature points

by using least squares algorithm, to ensure that the total error of the feature points after fine matching is less than 0.5 pixel. Finally, the geometry fine correction based on finely matched feature points is conducted.

```
┌─────────────────┐        ┌─────────────────┐
│  Images to be   │        │   Referential   │
│    matched      │        │     images      │
└────────┬────────┘        └────────┬────────┘
         ↓                          ↓
┌─────────────────┐        ┌─────────────────┐
│  Detection of   │        │  Detection of   │
│ feature points  │        │ feature points  │
└────────┬────────┘        └────────┬────────┘
         ↓                          ↓
┌─────────────────┐        ┌─────────────────┐
│ Description of  │        │ Description of  │
│  eigen vector   │        │  eigen vector   │
└─────────────────┘        └─────────────────┘
              ↓
    ┌───────────────────────┐
    │ Matching of feature    │
    │     point pairs        │
    └───────────┬───────────┘
                ↓
    ┌───────────────────────┐
    │ Deletion of error      │
    │   feature points       │
    └───────────┬───────────┘
                ↓
    ┌───────────────────────┐
    │ Further correction of  │
    │   feature point pairs  │
    └───────────┬───────────┘
                ↓
    ┌───────────────────────┐
    │ Estimation of          │
    │ transformation         │
    │   parameters           │
    └───────────┬───────────┘
                ↓
    ┌───────────────────────┐
    │      Resampling        │
    └───────────────────────┘

SIFT algorithm
```
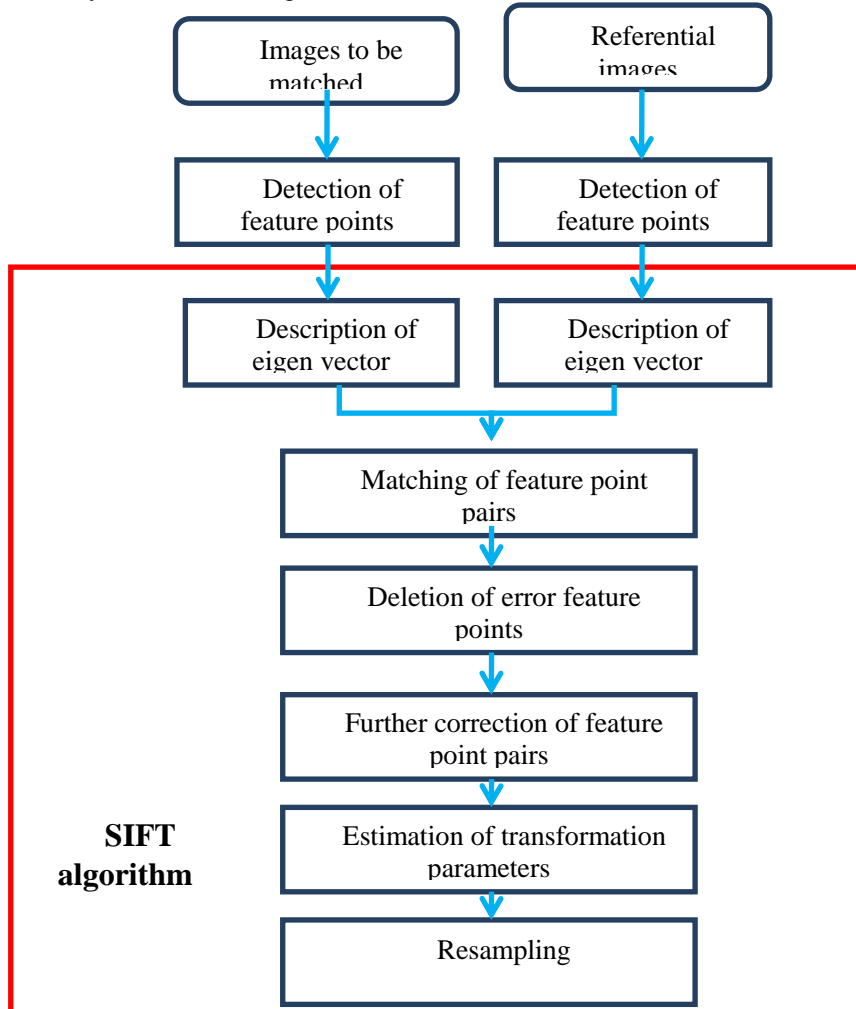
Figure 3. Flow Chart of geometry fine correction

Feature points and eigenvector refer to the local features of detected images (e.g. edge, angular point, outline, etc). Then it is necessary to combine and convert the features based on the need of matching targets, so as to form eigen vectors which are easy to be matched with good stability, to convert the image geometry correction issue into feature matching issue, and further to convert the feature matching issue into clustering issue of eigenvector of feature space. Eigen vector information is a set

of vectors with 128 dimensions. The following Figure illustrates part information of feature points and eigen vector of the Reference image.

The matching is conducted according to the similarity of eigen vector. Usually, various distance functions are taken as the similarity measurement of features, such as Euclidean distance and Mahalanobis distance, etc. The following figure is the schematic diagram of feature point location of two images after coarse matching. The SIFT algorithm is used to extract the 435 feature point pairs of two images.

A total of 50 pairs of feature points are randomly selected from 435 pairs of feature points and they are taken as the ground control points input into the ERDAS geometric correction procedure, to eventually get the error result, which is illustrated in the following Figure. In the figure, the error of the X direction is 4.9251, error of Y direction is 5.0334, and the total error is 7.0422.

GCP Tool : (Input : geomap.tif) (Reference : refmap.tif)

File View Edit Help

Control Point Error: (X) 4.9251 (Y) 5.0334 (Total) 7.0422

| Point #	Color	X Input	Y Input	>	Color	X Ref.	Y Ref.	Type	X Residual	Y Residual	RMS Error	Contrib.	Match
43		502232.174	4435316.708			1006532.947	4381497.111	Control	0.128	16.491	16.492	2.342	0.000
44		499852.815	4436153.281			1004118.181	4380954.893	Control	6.994	-8.002	10.628	1.509	0.000
45		500331.790	4435886.781			1004605.036	4381142.909	Control	-3.445	17.935	18.263	2.593	0.000
46		498202.702	4433953.825			1002769.108	4383376.247	Control	4.542	-5.193	6.899	0.980	0.000
47		499217.612	4433443.848			1003820.382	4383765.180	Control	0.754	1.145	1.371	0.195	0.000
48		500747.937	4434025.647			1005243.891	4382997.257	Control	-0.593	-1.625	1.729	0.246	0.000
49		501385.688	4434659.398			1005788.241	4382279.292	Control	-1.137	-0.466	1.229	0.175	0.000
50		500958.588	4435292.306			1005297.613	4381685.377	Control	3.031	2.611	4.001	0.568	0.000
51								Control					

Figure 4. Error result of coarse matching feature points as GCP

The fine matching on the feature points of the coarse matching is conducted by using least squares algorithm, and the final matching accuracy is optimized. The specific operation process is as follows:

- To establish the dualistic quadratic regression equation between the coarsely matched feature points.
- The feature points are calculated based on the regression equation of each feature point of image to be corrected.
- The error between simulation feature points and the feature points on Reference image is calculated and the maximum error is calculated.
- When the maximum errors are larger than the system preset default parameter of 0.5, the feature point is deleted and the procedure return to

step 1 to recalculate; otherwise, the fine matching of feature point pairs is completed. Then exit the calculation.

The following figure is the schematic diagram of feature point location after fine matching. The number of the feature point pairs after fine matching is 134.

Image to be corrected   Reference image
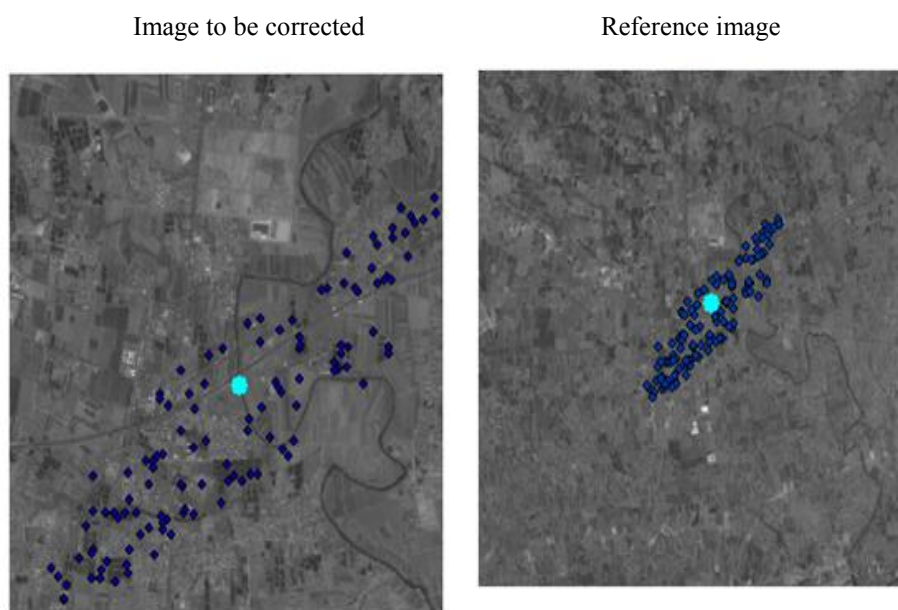
Figure 5. Schematic diagram of feature point pair location after fine matching

A total of 134 pairs of feature points are input into the ERDAS geometric correction module, to eventually get the error result, which is illustrated in the following Figure. The error of X direction is 2.4689, error of Y direction is 2.1418, and the total error is 3.2685.

Resampling of corrected image is conducted by using the feature points after fine matching, to finish the final geometry fine correction.

The following figure is the schematic diagram of the location of the image after fine correction and the original Reference image.

A total of 20 pairs of control points are randomly selected from the image after fine correction and the original Reference images, to conduct the accuracy analysis, and the final error result is obtained, which is illustrated in the following Figure.The error of the X direction is 1.1018, the error of Y direction is 0.8426, and the total error is 1.3870.

The paper analyses the limitation of the conventional SIFT algorithm in the matching of remote sensing images, and the extraction time and matching accuracy of feature points is optimized. Optimization algorithm can extract more matched feature points with higher accuracy. Based on the SIFT algorithm improved by this paper, and by firstly conducting coarse matching and then fine correction, the error is reduced from 7.0422 to 1.3870. The accuracy of the automatic matching of remote sensing images is improved, which is suitable for the business operation. However, the expenses for the system calculation time is relatively high. The further optimization on algorithm in the aspect of parallel computation could be conducted so as to reduce the operation time.

References

1. Zhang M Q. 2006. Research on the Line-Based Registration of Multi-Source Remote Sensing Imagery. Nanjing: Hohai University
2. Smith S M, Brady J M. SUSAN-A New Approach to Low Level Image Processing [J]. International Journal of Computer Vision, 1997, 23(1) : 45- 78
3. Harris C, Stephens M. A Combined Corner and Edge Detector[C]. Fourth Alvey Vision Conference. Manchester, UK, 1988:147-151
4. Xiangzeng Liu, Zheng Tian, Chunyan Chai, Huijing Fua. Multiscale registration of remote sensing image using robust SIFT features in Steerable-Domain[J]. 2011, 14(2): 63－72
5. Yves Dufournaud. Matching Images with Different Resolutions. Theory in Computer Vision[M] . Kluwer Academic Publishers, 2004
6. Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004,60(2) : 91—110.
7. Brown M, Lowe D G. Recognising Panoramas [A]. In Proceedings of the 9th
8. Schaffalitzky F, Zisserman A. Multi-view Matching for Unordered Image Sets, or How do I Organize my Holiday Snaps [A]Proceedings of the 7th European Conference on Computer Vision(ECCV02) [C]. 2002.
9. Lowe D G. Object Recognition from Local Scale-Invariant Features [C]. International Conference on Computer Vision , Corfu , Greece , 1999
10. Mikolajczyk K, Schmid C. A Performance Evaluation of Local Descriptors [J]. IEEE Trans Pattern Analysis and Machine Intelligence , 2005 , 27 (10) :161521 630
11. Li Q L, Wang G Y, Liu J G and Chen S B. Robust scale-invariant feature matching for remote sensing image registration[J], IEEE Geoscience and Remote Sensing Letters, 2009，6(2): 287–291