AUTO-MATCHING ALGORITHM FOR REMOTE SENSING IMAGES

H. C. Liu [a, b], G. J. He [a,*]

[a] Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, 100094, Beijing, China - (hcliu, gjhe)@ceode.ac.cn
[b] Graduate School of the Chinese Academy of Sciences, 100039, Beijing, China

KEY WORDS: Landsat-TM Image, Remote Sensing, Image Matching, Feature Matching

ABSTRACT:

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. The present differences between images are introduced due to different imaging conditions. As image registration geometrically aligns two images—the reference and non-corrected images, it becomes a crucial step in all image analysis tasks. Image registration is mainly divided into three steps: first is to choose matching points between the non-corrected image and the reference image, second is to determine the parameters and type of the mapping function with these matching points, third is to use mapping function to transform the non-corrected image. Among them, the selection of matching points usually adopts artificially selection manner whose efficiency is quite low. This paper will combine Harris algorithm, standardization cross-correlation algorithm, and least mean square algorithm to automatically match the Landsat-TM images. This method can produce matching points with high precision, and without manual work it also saves a lot of time.

1. INTRODUCTION

With the development of space technology, remote sensing has become an important data source. How to effectively use the information of the multi-band remote sensing images received from different sensors, different time to study environment and sustainable development process is an urgent new issue of earth observation [1]. In remote sensing image processing and analysis process, images have relative difference in geometry and radiation, which requires the registration of images prior to the analysis. Image registration is to align or broadly match two or multiple images from the same scene received by the same or different sensors (imaging conditions), in different conditions (weather, illumination, location and angle photography, etc).

Image matching is a crucial step in image registration. Existing automatic image matching methods can be divided into two main categories: region-based methods and feature-based methods. Region-based approach is often called template matching method [2], which combines the image feature detection and matching. Such methods do not try to detect a significant object in the image, but with a predefined size of the window image even the whole image to estimate the response. Fourier method and mutual information method belong to the region-based auto-matching method [3]. On the other side, the feature-based matching method is dependent on a significant image feature extraction. Obvious area (forest , lake , field), line (regional boundary, coastline, road, river) or point (region corner, line intersection) can be considered as feature. Alhichri and Kamel [4] proposed the idea of virtual circle, using distance transform method to extract invariant area features. Reference [5] used Harris operator in edge detection and corner points, and neighborhood affine invariant in the feature extraction. Line feature detection methods include the standard edge detection operators such as Canny operator, Gaussian Laplace and so on. Reference [6] introduced the existing edge detection algorithm and its evaluation. Li, et al [7] proposed a method which aims to detect the feature of the distorted image (with speckle noise of SAR data) on the basis of line features in the reference image (multi-spectral data). References [8-11] made a comprehensive and detailed introduction about both the classic and the latest corner detection algorithms, and especially in [11], the positioning characteristics of the detection operator were also analyzed. Since the corner is of good invariance when deformation occurred in the image, and it is easily perceived by the human eye, the corner is often used as control point. In addition, the method based on spatial relations [12], the constant descriptor methods, relaxation matching method [13] and the pyramid and wavelet methods are also part of feature-based image matching method.

Image registration process can be divided into three basic steps. First, select matching points between the distorted image and the reference image. Second, determine the parameters and types of the mapping function by the matching points. Third, transform the distorted image using the mapping function. But the manual selection and measurement of matching points is tedious, particularly in a production environment. In order to solve the problem, this paper introduced a new method of automatic matching points selection. The main idea of this method is: using the Harris corner detection operator to generate multiple feature points in the reference image, and determining a target window centered in each feature point. Then the template is moving in the corresponding search area in the distorted image, while the similarity comparison is calculated between the area covered by the template and the target window to find the location of maximum similarity. At last, the LMS algorithm is used for positioning the matching point, and ultimately determining the matching point pairs.

2. PRINCIPLES

2.1 Hirris Corner Detection Operator

One of the earliest corner detection algorithms is the Moravec's operator. The basic idea of it is that the corner can be easily recognized in a small window, because the shifting of the window in any direction should give a large change in intensity. The change of intensity for the shift *[u,v]* can be represented as following:

* Corresponding author.

$$\mathrm{E(u,v)}\Big|_{(x,y)} = \sum w(x,y)\left[I(x+u,y+v)-I(x,y)\right]^2$$

where $I\ (x+u,\ y+v)$ means the shifted intensity, $I(x,\ y)$ means the original intensity, and $w(x,\ y)$ means the window function where Gaussian function can usually be used.

Harris [14] improved Moravec's corner detector, by considering the differential of the corner score with respect to direction directly, instead of using shifted patches. It aims to find points with large corner response function, and takes the points of local maxima of this corner response. Average intensity change in direction *[u,v]* can be also expressed as a bilinear form (1):

$$E(u,v)\Big|_{(x,y)} = [u,v]M\begin{bmatrix} u \\ v \end{bmatrix} \quad (1)$$

$$M = \sum_{x,y} w(x,y)\begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix} \quad (2)$$

Here, I_x 、 I_y represent the derivative of the image in the horizontal, vertical direction. Similar diagonalization of the *M*:

$$M \to R^{-1}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R \quad (3)$$

In formula (3) , λ_1 and λ_2 are eigenvalues of *M*. So each pixel corresponds to such a quaternion matrix. As *R* can be seen as a rotation factor, the change of intensity can be analyzed only by the eigenvalues: when λ_1 and λ_2 both are small, defined as a flat area; when one large and one small, defined as the edge; when both are large, defined as a corner. Specific formula (4) is used to express this idea:

$$Cornerness = \det M - k(traceM)^2 \quad (4)$$

where $\det M = \lambda_1\lambda_2$, $traceM = \lambda_1 + \lambda_2$.

Cornerness means the measure of corner response, and *detM*, *traceM* respectively represent the determinant and trace of *M*. A good (corner) point should have a large intensity change in all directions, i.e. *Cornerness* should be large positive.

2.2 Normalized Cross Correlation（NCC）[15]

Normalized cross-correlation method uses statistical theory to automatically find matching points in the two images. An *n*-by-*n* “target” chip *T* (“template”) is selected in the reference image and an *m*-by-*m* “search” chip *S*, with *m* greater than *n*, is selected in the distorted image. A “template match” spatial similarity metric is calculated by sliding the target chip over the central *n*-by-*n* region of the search area (Fig. 1), multiplying the two arrays pixel-by-pixel, and summing the result for each shift location *(i,j)*.

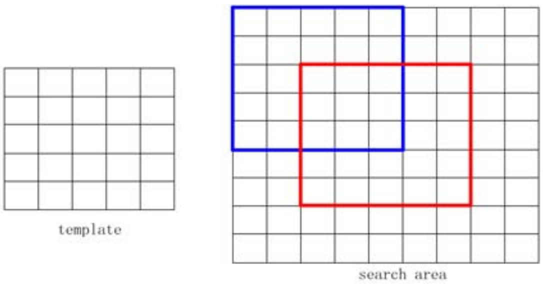

Figure 1. Example for template and search area
（the target or the template size is 5×5，where n=5, and the search area is 9×9，where m=n+4）

(a) reference image (b) distorted image

Figure 2. Example for NCC

The NCC used for finding matches of a template *t(x)* of size *n* in a signal *G(x)* of size *m* is defined as

$$r = \frac{\sum_{i=1}^{n1}\sum_{j=1}^{n2}\left(G_t(i,j)-\bar{G}_t\right)\left(G_s(i,j)-\bar{G}_s\right)}{\sqrt{\sum_{i=1}^{n1}\sum_{j=1}^{n2}\left(G_t(i,j)-\bar{G}_t\right)^2\sum_{i=1}^{n1}\sum_{j=1}^{n2}\left(G_s(i,j)-\bar{G}_s\right)^2}} \quad (5)$$

where the summations are over all template coordinates, $\bar{G}_t$ is the mean of the template and $\bar{G}_s$ is the mean of the region under the template in the other image. $n1, n2$ respectively means the size of the template and the region under the template in the other image.

2.3 Least Mean Square（LMS）[16]

LMS algorithm is used in this study to locate exact position of matching points between the template and the sifting window which has been found as the matching result (chip) in the distorted image. It aims to find the minimum of the square of intensity difference between the template and the chip by adjusting the geometry of radiation (affine transformation as an example). And then, the matching point in the chip need to be relocated in the transformed chip, as a result the new position of the matching point would be of higher accuracy.

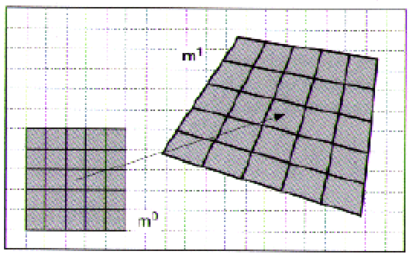

Figure 3. LMS illustration [Schenk,1999]

3. MAIN PROCESS

In this paper the automatic image matching method comprises Harris algorithm, NCC algorithm and LMS algorithm. Harris is used for generating multiple feature points in the reference image, NCC for area matching between the template and the shifting chip in the distorted image, and LMS for the exact location of the matching points. First, Harris corner detection operator generated multiple feature points in the reference image, and a target patch (size of 15 × 15) was determined centered in each feature point as the matching template T. And in the distorted image we defined a searching area according to the original geographical information which was larger than the size of the template. The DN arrays are correlated by "sliding" the target patch over the search area and calculating Eq. (5) at each possible shift position. After finding the most similar chip to the template, the LMS algorithm was used for matching point positioning, and ultimately determined the matching point pairs. Algorithm flow was shown in Fig.4.

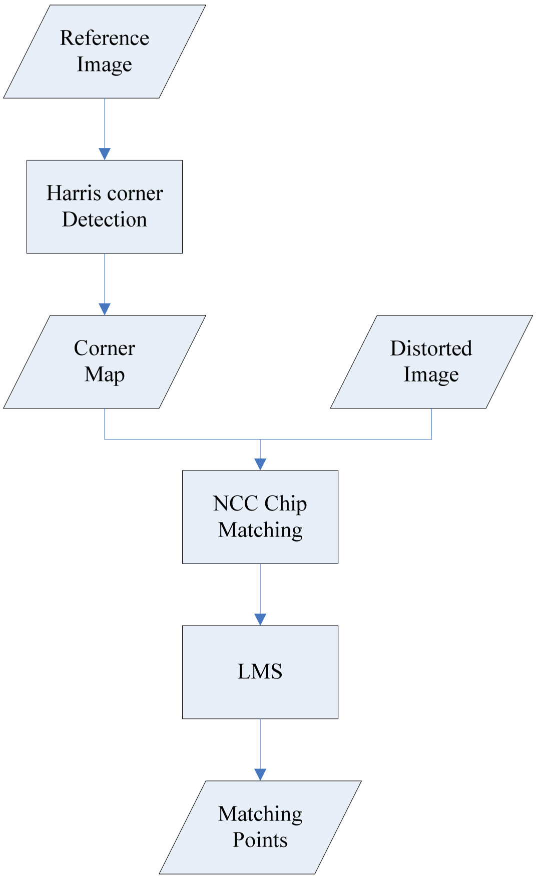

Figure 4. Flow chart

4. EXPERIMENT

This algorithm used a full scene Landsat-TM (119/042) image for experiment. The orthorected image data obtained on 15 June 1986, with TM projection, WGS84 ellipsoid, and spatial resolution of 28.5m, was selected as the reference image.

The distorted image data was the same scene obtained on 28 February 2008, systematically corrected with TM projection, WGS84 ellipsoid, and spatial resolution of 30m. This image has a initial positioning accuracy of 46 pixels, which is about 1380 meters.

Band4 of both images were used in the experiment, after the computation we got 532 matching points shown in Fig.5, and a more detailed representation in Fig.6.

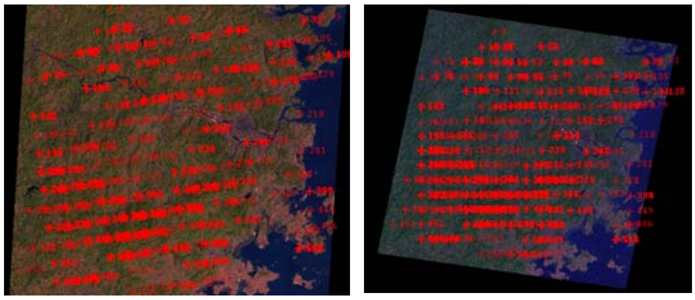

Figure 5. Matching result

(a) Distorted image (b) Reference image

Figure 6. Part of Fig.5

Furthermore, we also compared this result with the one obtained by auto-matching model in ERDAS software with the same experiment image data, shown in Tab.1.

Table 1. Result comparison

Methods	Num of points	X-Error (pixel)	Y-Error (pixel)	Error
Presented method	523	0.69	0.43	0.8130
Erdas auto-matching	350	0.78	0.35	0.8549

In contrast, our method and auto-matching in ERDAS had similar accuracy, but the former produced more matching points than the latter. At the same time, the format conversion from the original Landsat-TM to ".tiff" or ".img" is not necessary in our method but needed in ERDAS auto-matching, which is more efficient, and more suitable for processing distorted images in engineering batch mode.

5. CONCLUSION

The algorithm described in this paper was developed for the Landsat-TM image registration as the basis for further processing. It has been used in a number of related projects especially in the engineering situations, as the computer applied control action directly to the process without manual intervention. However, the downside is, if the image to be matched has the local deformations, the result in the selected matching points will have lower accuracy.

REFERENCES

[1] ZHANG Ji-xian, LI Guo-sheng, ZENG Yu. J. Remote Sensing, 2005, 9(1): 73.

[2] ZHANG Zhi-jia, ZHANG Yu, SHI Ze-lin, HUANG Bai-sha. J. Infrared and Laser Engineering, 2004, 33(6): 615.

[3] YANG Meng, PAN Quan, ZHANG Shao-wu, ZHU Ying, ZHAO Chun-hui, CHENG Yong-mei. J. Image and Graphics, 2010, 15(9):1376.

[4] H.S. Alhichri, M. Kamel. J. Pattern Recognition Letters, 2003, 24:1181.

[5] T. Tuytelaars, L.V. Gool. J. Computer Vision, 2003.

[6] D. Ziou, S. Tabbone. http://citeseer.nj.nec.com/ziou97edge.html, 1997.

[7] H. Li, B.S. Manjunath, S.K. Mitra. IEEE Transactions on Image Processing, 1995, 4:320.

[8] K. Rohr. J. Mathematical Imaging and Vision, 1994, 4:139.

[9] S.M. Smith. http://www.fmrib.ox.ac.uk/~spacesteve/susan.

[10] Z. Zheng, H. Wang, E.K. Teoh. J. Pattern Recognition Letters, 1999, 20:149.

[11] K. Rohr. Computational Imaging and Vision Series, vol. 21, Kluwer Academic Publishers, Dordrecht, 2001.

[12] Xu Wen-qing. J. Computer and Digital Engineering, 2009, 37(6) :136.

[13] JIANG Wan-shou, ZHENG Shun-yi, ZHANG Zu-xun, ZHANG Jian-qing. J. Geomatics and Information Science of WUHAN University 2003, 28(5) : 510.

[14] http://en.wikipedia.org/wiki/Corner_detection.

[15] J. P. Lewis. P. Vision Interface 1995 :120.

[16] T. Schenk. Digital Photogrammetry. Volume I, TerraScience, Laurelville, Ohio, 1999: 251.