

Feature-based image registration using the shape context

LEI HUANG*†‡, ZHEN LI†

†Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, Beijing, 100012, China

‡Graduate University of Chinese Academy of Sciences, Beijing, 100190, China

Image registration is a basic and important process for multi-sensor or multi-temporal remote sensing. In this letter, a new feature-based method named shape context is proposed for airborne multi-sensor image matching. This method has been found to be robust in hand-written digit and object recognition, and it is now introduced into remote sensing image matching after some adjustments. In the proposed method, control points (CPs) are extracted on the reference image, and edge features are extracted on the reference and the sensed image, respectively. The shape context exploits feature similarity between circular regions of the two images to find corresponding CPs on the sensed image. Finally, the sensed image is warped according to the CPs using thin-plate spline interpolation. This method is successfully applied to register airborne optical and multi-band synthetic aperture radar (SAR) images in two experiments, and the results demonstrate its robustness and accuracy.

1. Introduction

When disasters such as earthquakes and floods occur, airborne synthetic aperture radar (SAR) and optical images are efficient and flexible data sources that can provide repeated images within short time intervals. In key areas, large numbers of images are acquired; these images commonly need to be combined together to analyse the ground situation. Timely multi-sensor image registration can be a great challenge. Image registration is the process of overlaying two or more images of the same scene taken at different times, from different view points, or by different sensors (Zitová *et al.* 2003). In remote sensing, it is a basic process for many applications, such as change detection, fusion, mosaics (Chen *et al.* 2007) and Interferometric Synthetic Aperture Radar (InSAR). Traditionally, registration is performed with control points (CPs) that are selected manually on images. However, manual registration is a time consuming task, and many automatic methods have been presented to decrease labour intensity.

In general, automatic registration falls into two types: area- and feature-based techniques. Area-based methods are relatively accurate and are preferably applied when distinctive information is provided by grey values rather than by local shapes and structures (Zitová *et al.* 2003). Feature-based methods are typically applied when the local structural information is more significant than the image intensity information. They allow the registration of images of a completely different nature and can handle image distortions to some extent. Feature-based methods basically consist of four steps: feature extraction, feature matching, transform model estimation and image transformation. In this letter, attention is concentrated on feature matching. In previous research, many matching strategies have been employed for remote sensing images. The Hausdorff distance measures the extent to which each point of a data set lies near some point of an 'image' set and *vice versa* (Huttenlocher *et al.* 1993), and it is improved to reduce the computational complexity for use in remote sensing (Mount

* Corresponding author. Email address: hlhjsx@126.com

1999). Patch features are used in SPOT and ERS-1 image matching, and the area, perimeter length, length and width are used to determine the similarity between patches (Dare *et al.* 2001). In urban areas, line features can be extracted from roads and be matched with the Modified Iterated Hough Transform (Habib *et al.* 2005). Invariant-moment shape features are extracted and chain-code matching is used (Dai *et al.* 1999). In Wen *et al.* (2008), spatial relations and organic feature similarity are combined as a matrix, and its global maximum is assumed to be reached when two images match well with each other.

Generally, the feature matching strategies mentioned above rely on “strong” features, such as special points, islands, closed lakes or straight lines. A broadly suitable algorithm called shape context is introduced, even for “weak” feature areas. As far as the authors know, this study is the first time that shape context has been used for remote sensing images. The study areas consist of flat farmland, in the Sichuan Province of China, and the usual feature is the field ridge. The data are obtained from airborne optical and multi-band SAR images.

2. Method

The shape context is presented for object recognition on silhouette images (Belongie *et al.* 2002), and it is briefly introduced in this section. In addition, some adjustments are made to make it fit for use with complex remote sensing images.

2.1 Shape Context

Edge features extracted from images are considered as point sets. For a point on the first image, it is expected to find the best matching point on the second image according to the edge features. It is identified that the distribution over relative positions is a more robust and compact, yet highly discriminative descriptor. For a point p_i on the shape, a coarse histogram h_i of the relative coordinates of the remaining $n-1$ points is computed:

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\} \quad (1)$$

(Figure 1)

This histogram is defined as the shape context of p_i . The bins that are uniform in log-polar space make the descriptor more sensitive to the positions of nearby sample points than to those of points further away. An example is shown in figure 1(c).

Consider a point p_i on the first shape and a point q_j on the second shape. Let $c_{i,j} = C(p_i, q_j)$ denote the cost of matching these two points. As shape contexts are distributions represented as histograms, it is natural to use the χ^2 test statistic:

$$c_{i,j} \equiv C(p_i, q_j) = \frac{1}{2} \sum_1^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (2)$$

where $h_i(k)$ and $h_j(k)$ denote the K -bin normalised histograms at p_i and q_j , respectively; c_{ij} is the cost to match two point sets.

2.2 Adjustments

The radius and orientation of the circular template are adjusted to make the algorithm adequate for remote sensing image matching. According to the definition of log-polar bins, pixels are indexed by the ring number R and the wedge number W . The radii of the rings surrounding the centre are $r, 2r, 4r, 8r,$ and $16r$. If r is set to two pixels as the radius of the smallest ring, the result becomes unstable, because the radius is too small to tolerate SAR image distortions. Therefore it is suggested that r equals four pixels.

The radius is also related to the scaling. The spatial resolutions of a reference and a sensed image are known, and their ratio is λ . λ can be used to adjust the circular template size when the two images have different spatial resolutions. If the radius used in the circular template is r for a reference image, then the radius should be λr for the corresponding sensed image.

The largest adjustment is that the circular template is given an orientation. In silhouette image matching, one can use a relative frame, based on treating the tangent vector at each point as the positive x -axis. For multi-sensor images, this does not work because edge features are too complex to stably determine the tangent vector. In a more practical way, it is supposed that the SAR image has a rotation, compared to an ortho-image; thus, the circular template is given a general rotation angle α to correspond with the SAR image, as shown in figure 2.

(Figure 2)

Compared to the spaceborne data, the swath of an airborne SAR image is much narrower. In addition, the study areas are flat farmland, so the distortion arising from terrain can be neglected. The side-looking airborne radar image is formed along the flight line of an airplane, and the image rotation angle α can be estimated from the flight line. In figure 3, A is the starting point and B is the end point of the flight line. The image rotation angle can be estimated from equation (3):

$$\alpha = \arctan \frac{Y_B - Y_A}{X_B - X_A} \quad (3)$$

X_A, Y_A, X_B, Y_B are the coordinates of A and B, obtained from a differential global positioning system (DGPS) receiver. The flight line is designed to be a straight line that measures several kilometres.

(Figure 3)

The image rotation angle obtained from the flight line is only an estimated value, and it may be inaccurate. However, it is confirmed that the shape context method is robust enough to tolerate small angle deviations. For an image I, the image is rotated through an angle θ , forming a new image I', as in figure 4.

(Figure 4)

To register the images I and I', a rotation angle $\theta + \varepsilon$ is assigned to the circular template, where ε is the given angle deviation. Registration is performed when ε increases from -20° to $+20^\circ$. As shown in figure 5, when $-3^\circ \leq \varepsilon \leq 3^\circ$, the average coordinate deviation is within 0.4 pixels. In most cases, the angle deviation ε is between -1° and 1° , which translates to a deviation of about 0.2 pixels, according to figure 5. The only difference between images I and I' is the rotation angle θ , so the coordinate deviation mentioned above arises from the

angle deviation ε . The experiment demonstrates that the rotation angle estimated from the flight line is sufficient to obtain accurate registration results using the shape context. (Figure 5)

2.3 Matching Procedure

The entire matching procedure was performed using the following steps.

1) Special preparations for the SAR images. If the registration is performed on optical and SAR images, the SAR image should be converted from slant range to ground range. All of the employed SAR images are smoothed with an enhanced-Lee filter that is especially designed to suppress speckle noise (Lee 1981).

2) Control points $M = \{m_i\}$ are extracted from the reference image with the Harris operator (Harris *et al.* 1988).

3) Edge features are extracted from both the reference and the sensed images with the Canny operator (John 1986).

4) For each control point m_i on the reference image, the local edge feature distribution is measured based on the shape context. Assuming that the candidate matching point is within an $N \times N$ pixel area on the sensed image, the matching cost is calculated between m_i and each point in the area with equation (2). A lower $C_{i,j}$ represents a higher similarity, and the point with the lowest $C_{i,j}$ is selected as the corresponding point of m_i . The selected points in the sensed image constitute a set of CPs $S = \{s_i\}$.

5) The invalid CPs are removed in this step. The three points that take the lowest cost in shape context matching are selected to estimate the affine transformation parameters in equation (4)

$$\begin{aligned} x &= a1x'+b1y'+c1 \\ y &= a2x'+b2y'+c2 \end{aligned} \quad (4)$$

where $a1, b1, c1, a2, b2, c2$ are affine transformation parameters, and (x', y') and (x, y) are corresponding CPs in the reference image and the sensed image. Invalid CPs are removed by comparing the distance between the coordinates of the mapped control points and the actual points.

6) Thin-plate splines (TPS) are used to warp the sensed image. These are perhaps the most widely used transformation functions in image registration with nonlinear geometric differences. The warping of the sensed image with respect to the reference image is accomplished with equation (5), where (x', y') and (x, y) are all of the valid CPs selected in step (5); w_i is the weight of the nonlinear radial interpolation function K . The details of TPS are introduced in Bentoutou *et al.* (2005).

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^n W_{xi} K(\|(x, y) - (x_i, y_i)\|) \\ \sum_{i=1}^n W_{xi} K(\|(x, y) - (x_i, y_i)\|) \end{pmatrix} \quad (5)$$

3. Experiments and analysis

To evaluate the robustness of the proposed registration method, two experiments were performed. First, the method was performed for airborne optical and SAR images; second, it was then also performed for multi-band SAR images.

3.1 Registration of Airborne Optical and SAR images

The original slant range SAR image is converted to a ground range image. Being a ranging device, radar records objects according to the distance from the aircraft to the object, thus, forming a slant range image. By applying a correction from slant range to ground range, the scale relationship between the image and the ground becomes linear. This is only an approximation and is based on the assumption of level topography (Henderson *et al.* 1998). The pixel sizes in the azimuth and the range directions of the SAR data are both 0.5 m, which accord with the optical image.

(Figure 6)

The airborne optical ortho-image, obtained from ADS40, and the C-band SAR image are employed in the first group of experiments. The rotation angle estimated from the flight line is 7.45° and is used to rotate the circular template. The rotation angle calculated from CPs using the Cartesian coordinate system (Goshtasby *et al.* 1986) is 6.96° , which is close to the estimated value. The employed data and the matching results are shown in figure 6.

3.2 Registration of Multi-band SAR Images

Airborne L-band and C-band SAR images are employed in the second group of experiments. The two images arise from parallel flight lines (perpendicular distance of about 200 m), so the estimated rotation angle between the two images is zero. The spatial resolutions of the two kinds of SAR data are the same. L-band radar operates with a longer wavelength than C-band radar. The different penetration of the radar signals into the ground and plants results in different image textures. The employed data and matching results are shown in figure 7. The registration is performed on slant range images directly because the two images possess similar distortions.

(Figure 7)

4. Accuracy Analysis.

To estimate the accuracy of the final registration results using the proposed method, all of the valid CPs were used in the evaluation. Parts of the CPs are listed in table 1. ΔX and ΔY are absolute deviations of CPs between the reference image and the warped image. The root mean square error (RMSE) involving all of the CPs in experiment A is 1.883 pixels in the X direction and 1.752 pixels in the Y direction, while in experiment B they are 1.138 and 0.964 pixels. The optical and SAR images have a lower matching accuracy than the multi-band SAR images. As mentioned above, multi-band SAR data arise from parallel flight lines, so they have similar distortions and are more accurately registered.

(Table 1)

5. Conclusions

This letter develops a new method called the shape context, which matches images by comparing edge-feature distributions in a defined circular template. In this study it is used on remote sensing images for the first time after some practical adjustments. The shape context method is invariant to rotation and scale when matching the boundaries of silhouette images. However, it is hard to implement it on remote sensing images, which are much more complex. For the employed airborne SAR data, the rotation angle was estimated according to the flight line, and the scale ratio was supposed to be known; thus, the image registration was performed without any manually selected CPs. The proposed method was validated on optical-SAR images and multi-band SAR images in the experiments. It is confirmed to be an efficient and broadly suitable feature-based method that can be used in multi-sensor registration. In future work, the accuracy and feasibility of the proposed method will be tested on satellite sensor data using orbit parameters.

Acknowledgements

Our research was supported by the Chinese Ministry of Science and Technology (Grant No. 2009CB723901, 2009AA12Z122, 2009AA12Z145).

References

- BELONGIE, S., MALIK, J., Puzicha, J., 2002, Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, pp. 509-522.
- BENTOUTOU, Y., TALEB, N., KPALMA, K., and RONSIN, J., 2005, An automatic image registration for applications in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 43, pp. 2127-2137.
- CHEN, F., WANG, C., ZHANG, H., 2007, Automatic matching of high-resolution SAR images. *International Journal of Remote Sensing*, 28, pp. 3665-3678.
- DAI, X., KHORRAM, S., 1999, A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Transactions on Geoscience and Remote Sensing*, 37, pp.2351-2362.
- DARE, P., DOWMAN, I., 2001, An improved model for automatic feature-based registration of SAR and SPOT images. *Journal of Photogrammetry & Remote Sensing*, 56, pp. 13-28.
- GOSHTASBY, A., STOCKMAN, G.C., PAGE, C.V., 1986, A region-based approach to digital image registration with subpixel accuracy. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24, pp. 390-399.
- HABIB, A., AL-RUZOUQ, R., 2005, Semi-automatic registration of multi-source satellite imagery with varying geometric resolutions. *Photogrammetric Engineering & Remote Sensing*, 71, pp. 325-332.
- HARRIS, C., STEPHENS, M., 1988, A combined corner and edge detector. In *Fourth Alvey Vision Conference*, UK, pp. 147-151.
- HENDERSON, F.M., LEVIS, A.J.(Ed.), 1998, *Principles and Applications of Imaging Radar, Manual of Remote Sensing*, 3rd edition, vol 2. pp. 142-149 (New York: Wiley).
- HUTTENLOCHER, D.P., G.A.Klanderma, W.A.Rucklidge, 1993, Comparing images using the Hausdorff Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15, pp. 850-863.
- JOHN, C., 1986, A computational approach to edge detection. *IEEE Transactions on Geoscience and Remote Sensing*, PAMI-8, pp. 679-698.
- LEE, J.S., 1981, Speckle analysis and smoothing of synthetic aperture radar images. *Computer Graphics and Image Processing*, 17, pp. 24-32.
- MOUNT, D.M., NETANYAHU, N.S., MOIGNE, J.L., 1999, Efficient algorithms for robust feature matching. *Pattern Recognition*, 32, pp. 17-38.

WEN, G.J., LV. J., YU. W., 2008, A high performance feature-matching method for image registration by combining spatial and similarity information. *IEEE Transactions on Geoscience and Remote Sensing*, 46, pp.1266-1277.

ZITOVÁ, B., FLUSSER, J., 2003, Image registration methods:a survey. *Image and Vision Computing*, 21, pp. 977-1000.

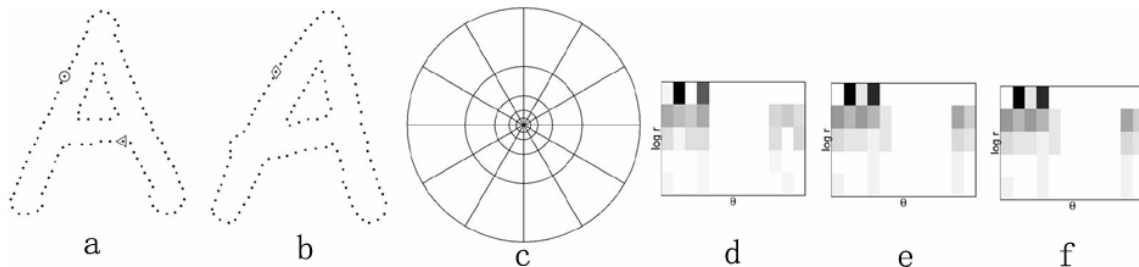


Figure 1. Shape context computation and matching. (a) and (b) Sampled edge points of two shapes. (c) Diagram of log-polar histogram bins used to compute the shape contexts. Five bins for $\log r$ and 12 bins for θ are used. In the following text it is called the circular template. (d), (e), (f) Example shape contexts for reference samples marked by \circ , \diamond , and \triangleleft in (a) and (b). Each shape context is a log-polar histogram of the coordinates of the rest of the point set measured using the reference point as the origin. (Dark=large value) Note the visual similarity of the shape contexts for \circ and \diamond which were computed for relatively similar points on the two shapes. In contrast, the shape context for \triangleleft is quite different.

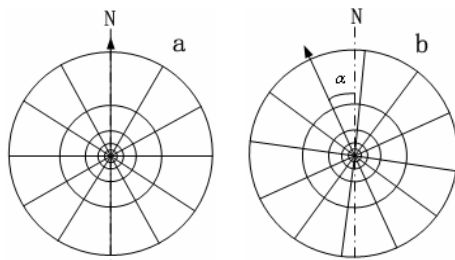


Figure 2. Rotation of histogram bins. (a) Circular template without rotation. (b) Circular template with rotation by angle α . The rotation of the circular template is in accord with the flight line.

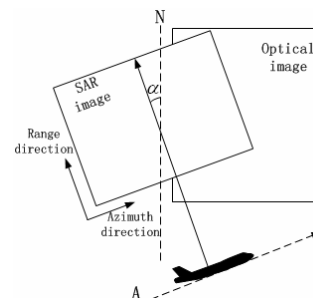


Figure 3. Estimation of rotation angle from the flight line of the airplane

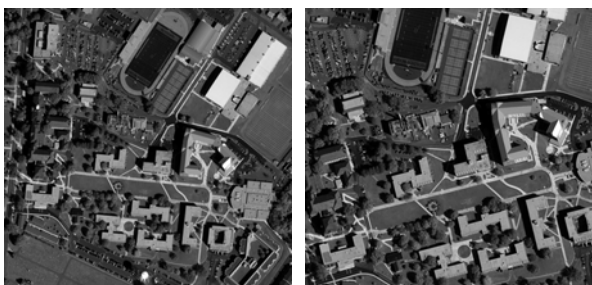


Figure 4. Left image I, a GeoEye image downloaded from the internet; on the right is image I' which is rotated through an angle $\theta=20^\circ$ from image I.

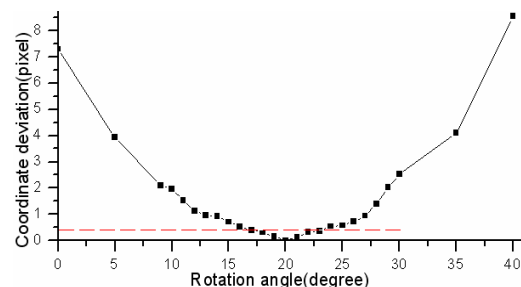


Figure 5. Curve diagram of rotation angle matching accuracy. A total of 76 points are matched on images I and I' of figure 4. The averaged deviation of the 76 points is calculated when the rotation angle of the circular template changes from 0-40.

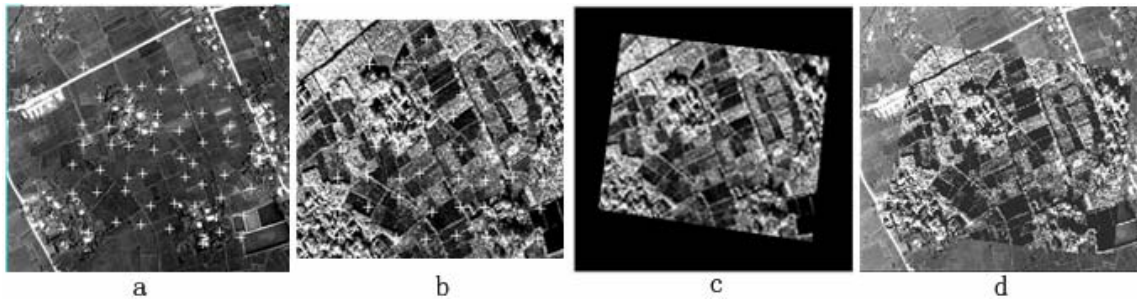


Figure 6. Registration of optical and SAR images. (a) CPs on the reference image. (b) CPs on the sensed image (c) Warped sensed image. (d) The matching result.

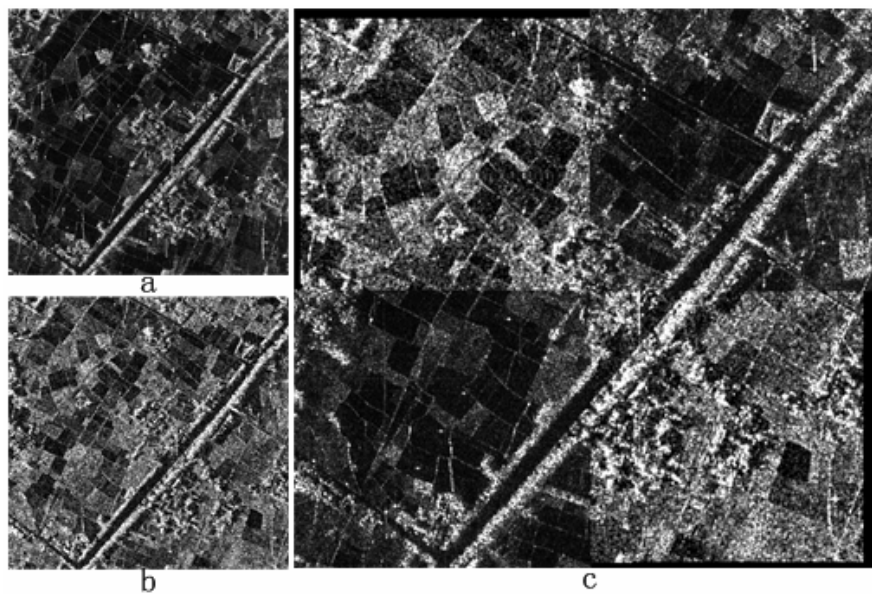


Figure 7. Registration of multi-band SAR images. (a) L-band SAR image. (b) C-band SAR image. (c) Matching result of the L and C bands. Top left and bottom right of image c are from the C-band image, and the other two parts are from the L-band image.

Table 1. Results of the accuracy evaluation.

Point	Experiment A		Experiment B	
	ΔX	ΔY	ΔX	ΔY
1	1.933	1.692	1.992	0.48
2	2.108	0.819	0.415	0.2
3	1.195	3.048	0.452	1.646
4	1.078	1.901	0.36265	1.405
5	1.216	1.574	1.44	0.265
6	0.376	0.657	2.395	0.068
7	3.073	0.628	0.03492	0.791
8	0.873	0.261	1.001	1.731
9	2.859	3.353	0.203	0.48
10	3.92	2.01	0.45	2.24
RMSE	1.883	1.752	1.138	0.964