

# Object-Based Multi-mode SAR Image Matching

Jie Rui<sup>1,2,3(✉)</sup>, Chao Wang<sup>1</sup>, Hong Zhang<sup>1</sup>, Bo Zhang<sup>1</sup>, Fan Wang<sup>3</sup>,  
and Fei Jin<sup>3</sup>

<sup>1</sup> Institute of Remote Sensing and Digital Earth,  
Chinese Academy of Sciences, Beijing, China  
ruijie@radi.ac.cn

<sup>2</sup> University of Chinese Academy of Sciences, Beijing, China

<sup>3</sup> Zhengzhou Institute of Surveying and Mapping, Zhengzhou, China

**Abstract.** Owing to the effect of imaging mechanism and imaging conditions in synthetic aperture radar (SAR) image, inconsistent features and relationship correspondence constitute key problems using traditional image matching algorithms because of significant differences between the images. This study proposes an object-based SAR image matching method. Two images are matched through same ground objects, by means of property and shape information of objects, which are obtained via object extraction and morphological operations. We utilize a shape context descriptor to compare contours of objects and detected invariant control points. The experimental results show that the proposed method achieves reliable and stable matching performance, and can alleviate deformation and nonlinear distortion effects of different systems.

**Keywords:** Image matching · Object-based · Shape matching · Shape context

## 1 Introduction

Image matching technology is an extremely important modern information processing technology, especially in the field of image processing, to integrate multisource and multitemporal data. It has been widely used in multispectral classification, object recognition, image fusion, change detection, image mosaicking, image sequence analysis, and cartography updating.

Image matching focuses on searching similar objects or scenes in mass images. It is different to image registration, which focuses on geometrically aligning two images of the same scene. Most methods of both image matching and image registration are same because both need to find similar features in images, except the differences of emphasis on search strategy and geometric precision. Image matching is considered to be pre-processing or preliminary work for image registration in most applications, primarily because relationship of features or objects between different images is difficult to establish, which is usually referred to as correspondence problem in literatures. Although a lot of work has been done, automatic image matching still remains an open problem. Until now, it is still rare to find a rapid, robust, and automatic image matching method, and most existing image matching methods are designed for particular

applications. The performance of a methodology is always limited to a specific application, or sensor characteristics and the terrain characteristic of the imaged area [1].

## 2 Image Matching

Image matching methods can be broadly classified into two main categories: area-based and feature-based. Area-based methods measure region correlation directly using information of image intensities, e.g., correlation coefficient (CC), mutual information (MI), without any structural analysis. Consequently, they are sensitive to the intensity changes and have large amount of data to be processed. The limitations are optimal size of window and nonlinear transformation between two images. Feature-based methods do not consider the pixel grey values as a primitive to correlate, but extract other structural information (e.g. points, edges, linear structures, regions) [2, 3]. The feature-based methods have relative small computation, and the distinguished advantages are high reliability and speed, which are more appropriate for areas with poor information. The drawback of the feature-based methods is that the respective features might be hard to detect and/or unstable in time.

Sophisticated approaches have been developed based on multi-features or multi-layer feature [4]. Meanwhile, matching methods using both area-based and feature-based approaches simultaneously are received more attention [5, 6]. The integrated methods mainly include two steps: roughly match images based on feature extraction, and further improve reliability and precision based on correlation-like method. Moreover, some optimization approaches (e.g. Genetic algorithm, colony optimization) and matching strategies (e.g. tree/graph matching) also be used in matching process [7, 8].

The significant differences between the multisource images and most of the inherent speckle noise in the SAR images appear when matching multi-mode SAR images due to different imaging mechanisms and imaging conditions. For multisource imagery, the texture and grey levels at the location of conjugate positions will not likely be similar. Consequently, relationships between extracted points or lines features in different images will be difficult to establish, which cannot be solved perfectly through the sophisticated methods or optimization approaches.

The features extracted automatically using computer vision techniques are mostly low-level features (e.g. point, edge, color, texture), which is coarse understanding of image scenes. Although low-level features can accurately be extracted by various methods, they cannot readily be utilized to describe human visual perception of an image. Human beings tend to use high-level features, such as objects, scenes, to interpret images and measure their similarity. Objects and scenes are real understanding of image semantics [9]. For remote sensing images, semantic perception focusses on objects rather than scenes in most applications. At present, a lot of research is about the object-based methods in computer vision and pattern recognition, e.g., content-based image retrieval (CBIR) [10, 11]. In the field of remote sensing, image segmentation, classification, and change detection using object-based methods has attracted much attention [12]. Very little research has explored the topic in image matching. Object extraction using segmentation technology and shape matching between complex and incomplete contours of objects are key problems. Some earlier literatures [13, 14]

exploit special objects (e.g. water regions, roads) for image matching, however, the purpose is only to detect distinct points (e.g., corners, centroids, intersections) as control point candidates. References [15, 16] exploit homogeneous closed region for multisource image matching and only edge or shape feature is used. In other words, these methods do not use semantic features of images and still belong to conventional feature-based methods.

In this paper, a novel object-based matching method is proposed. Obvious ground objects are extracted and two images are matched based on same objects or scenes by means of property and shape information of these objects, which belongs to high-level semantic features of images. We are interested in the application of object-based method and focus on robust shape matching algorithm. Curve smoothing, shape matching, and other related technology are used in matching process. The method has reliable and stable matching performance, and can alleviate the effect of deformation (rotation, translation, and scale) and nonlinear distortion caused by different sensors and different incident angles.

### 3 Proposed Method

#### 3.1 Rationale

The idea of using ground objects for matching sensed images originally came from human visual perception where fast recognition between two images is compared via obvious ground objects. These ground objects are processed as integral object, and subsequently, properties, geometry information, and relations between different objects are combined and comprehensively utilized. Contours of objects in different images are not completely coincident and have local difference due to various reasons, especially in multisource imagery. It is not a problem since similar objects can be matched by its main body (Fig. 1a). In contrast, meaningless points or lines are very difficult to establish the relationship, especially in the case of different number (Fig. 1b). Obviously, a great number of points or lines may be a more difficult problem, which is so-called correspondence problem mentioned previously.



**Fig. 1.** (a) Objects are easy to match even with local deformations. (b) Points (centroids of lakes) are difficult to match.

The key advantage of the proposed method is that extracted objects include some properties (e.g., category, contour length, size, etc.) and these properties are also used for matching. This is main difference, also important difference, from the methods based on simple shape information. The ground objects, as high-level semantic features, are real understanding of images. In this sense, the proposed method can be considered as high-level image processing based on image interpretation, accords with human visual perception. Thus, stability and reliability of object-based matching method exceed the traditional methods based on meaningless point, line and edge features.

Compared with point or line features, the object-based matching method has distinct advantages. Information content of objects is huge and information content is a well-known criterion to judge about the feature distinctiveness. The most important point is that correspondence problem can be easily solved by conjugate object pairs. Moreover, the number of objects extracted from images is very small and the search space is restricted to same category, e.g., lake-to-lake, vegetation-to-vegetation. It can rapidly reduces the computation of searching and avoids false matching between similar shapes from different category, e.g. lake and vegetation.

### 3.2 Matching Process

The proposed matching method is divided into three steps.

- (1) Objects extraction. First, the interest ground objects, e.g., lake, river, road, village, and vegetation, are extracted from the reference and the sensed image, respectively. The objects with different category use different segmentation algorithms. The extracted objects are marked with its category. Morphological Processing and data vectorization should be used to get shapes of objects.
- (2) Object matching. All properties (e.g., category, contour length, size) are used for object coarse matching. Shape processing is essential for shape matching and curve smoothing can reduce the difference of object shapes in different images. Then, object shapes between two datasets are compared one by one based on shape context descriptor. Of course, the shape matching is limited to same category, e.g., lake-to-lake, vegetation-to-vegetation. The two shapes matched successfully are conjugate object pairs.
- (3) Geometric transformation. Control points are selected from conjugate object pairs and the transformation estimation is solved via the least squares technique based on these control points. Matching precision is evaluated by computing residual error of check points.

The flowchart of the proposed method is presented in Fig. 2.

### 3.3 Shape Matching Using Shape Context

Shape is an important visual feature of an image, which provide most easy and obvious recognition information. However, shape matching is still a challenging task, especially, the objects in the sensed images often have a complex irregular shapes. The key

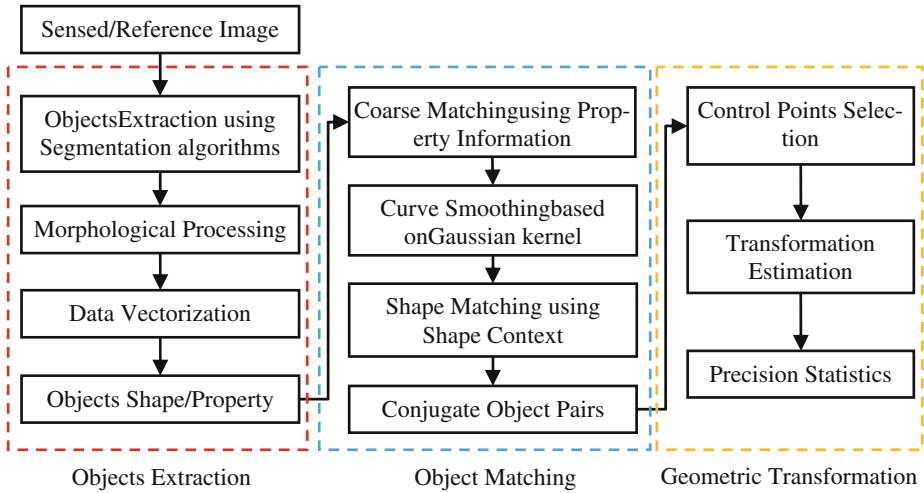


Fig. 2. Image matching flowchart.

point is to develop an appropriate descriptor to characterize the object shape and tolerate local deformations simultaneously. Therefore, the robust matching method for variform objects in remote sensing images is required.

Although numerous shape representation and analysis methods based on contour information exist, the most suitable tool to use for object contours is the shape context [17]. The shape context at a reference point captures the distribution of remaining points relative to it, thus generating a globally discriminative characterization. The shape context uses a log-polar histogram of coordinates and is tolerant of local deformations. The shape context can be used to quickly match a search for similar shapes via recognition.

For a point  $p_i$  on the shape, the histogram  $h_i$  is defined to be the shape context as follows

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

Where  $k \in \{1, 2, \dots, K\}$  and  $K$  is the number of histogram bins. We use 5 bins for polar radius and 12 bins for polar angle.  $(q - p_i) \in bin(k)$  means that the vector originating from  $p_i$  to another point  $q$  belongs to the  $k$ th bin.

Note visual similarities among shape contexts shown in Fig. 3. Corresponding points on two same shapes will have similar shape contexts (Fig. 3a and c), enabling us to solve for correspondences as an optimal assignment problem, also can be used for shape similarity measurement.

The shape context descriptor is based on the characteristics of histogram, and the  $\chi^2$  test statistics is used to measure the difference of two points. The cost of matching two points is defined as

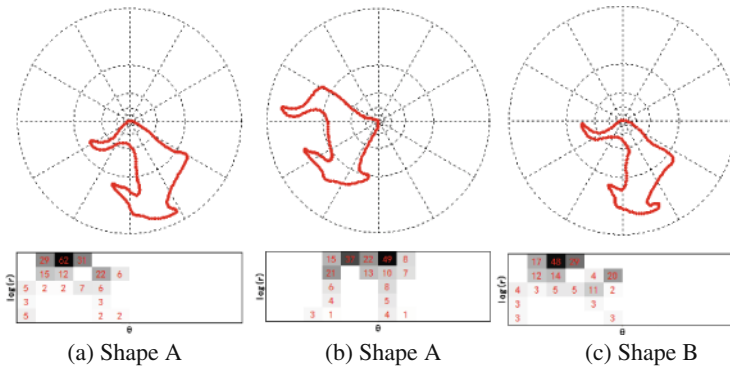


Fig. 3. Sampled edge points of two contours and example histogram of shape contexts.

$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \tag{2}$$

where  $p_i$  is a point on the first shape and  $q_j$  is a point on the second shape.  $h_i(k)$  and  $h_j(k)$  denote the K-bin normalized histogram at  $p_i$  and  $q_j$ , respectively.

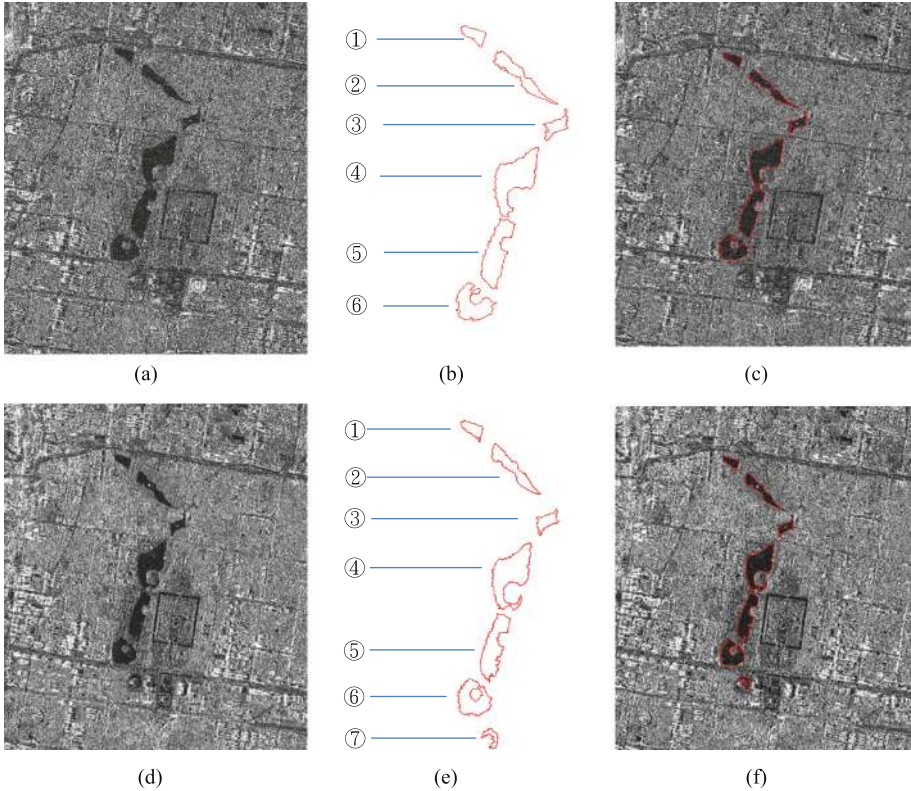
For  $p_i$ , traversing all points of the second shape and computing the point-to-point cost, the minimum value (which is less than the threshold) corresponding to the point is the matched point. Similarly, we can get the matching point of  $q_j$ . Two points are best paired if they correspond to one another.

In this case, the shape context descriptor is only used to match conjugate pairs of object contours between the two datasets, i.e., to find similar contours, regardless of correspondence between the points of the two shapes. In the design of our algorithm, two contours to determine whether similar based on matched point, that is to say that most are paired case, for example, more than 2/3, can be considered that the two contours are similar. While conventional algorithms compute shape similarities based on the shape context, such computations are complex and the threshold of similarity is difficult to determine. The proposed algorithm is simple to compute and presents a suitable similarity threshold, as few contours in different objects are similar.

### 4 Experimental Results

Two SAR satellite remote sensing image in Beijing, China, are used for our experiment. Radarsat-2 image, acquired in 2008 with the spatial resolution of 1.5 m, is regarded as the reference image. TerraSAR image, acquired in 2011 with the spatial resolution of 1.25 m, is regarded as the sensed image. Features in urban area from multi-mode SAR images are obviously inconsistent and relationship of features is difficult to established, indicates that the match will be a challenge.

In this experiment, we only select water region as the matching object since there are few constant distinct objects in urban area. Appropriate merge strategy is used after

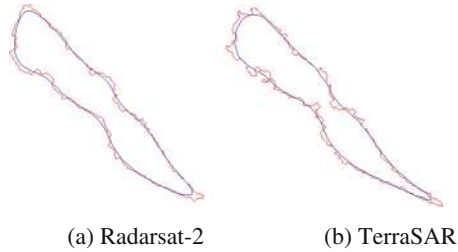


**Fig. 4.** SAR images in Beijing, China, and extracted objects. (a) Radarsat-2, 2008, (b) extracted result and (c) contours overlay on the original image (a). (d) TerraSAR, 2011, (e) extracted result and (f) contours overlay on the original image (d).

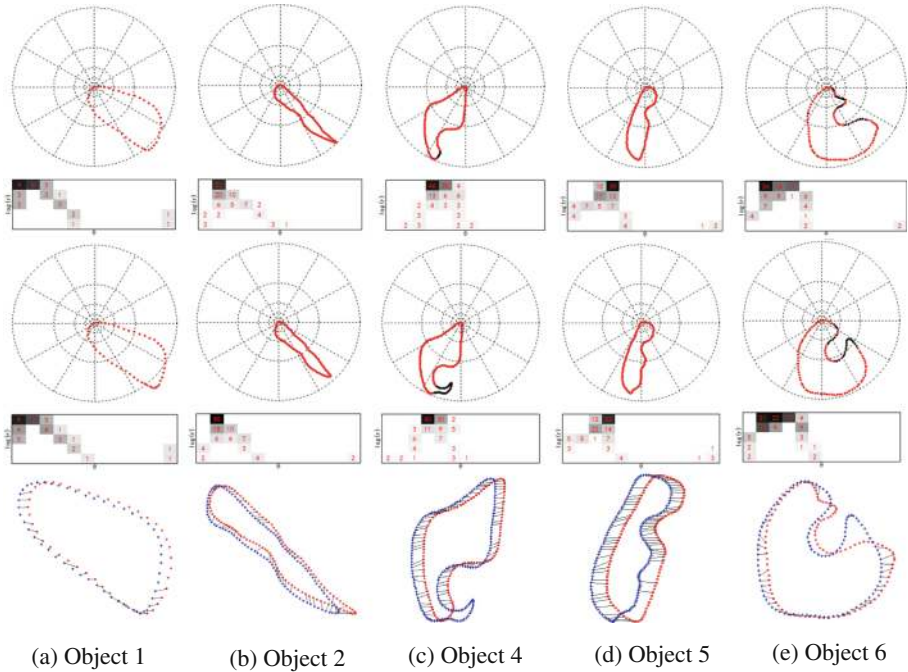
segmentation for whole objects. Two images with extracted objects are shown in Fig. 4. It can be seen that shapes of same object have distinct local difference.

The contours of objects extracted using segmentation and vectorization algorithms are quite coarse and have a lot of burrs. Thus, curve smoothing is essential to remove these burrs and details, i.e., reduce the difference of object shapes, which is useful for shape matching. An example of Gaussian smoothing is shown in Fig. 5.

We select a contour of the sensed objects and compute the maximum distance between all points. This distance is as the radius of the shape context descriptor, and the shape context is computed in accordance with a uniform radius for all contours of the reference objects. We match shapes between this object and all reference really objects individually and find optimal or suboptimal matching results (i.e., a similar shape cannot be found). Every contour of the sensed objects has a different radius during matching processing, and histogram distributions and the costs of all reference contour points are re-evaluated for each sensed object. The diagrams of object matching using shape context descriptor are shown in Fig. 6.

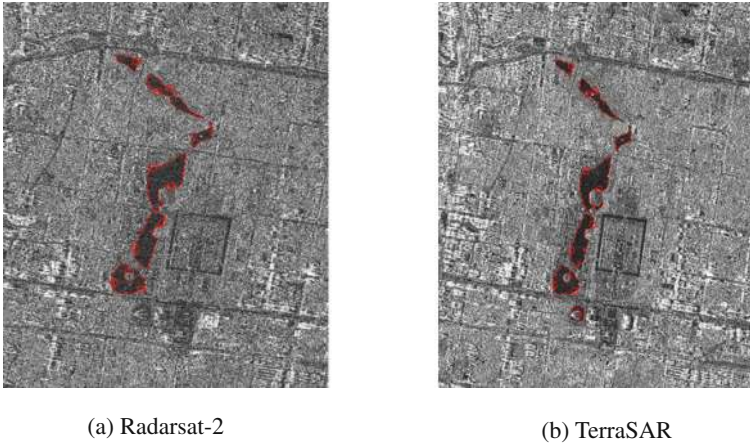


**Fig. 5.** Curve smoothing of Object 2. Red line is original extracted contour and blue line is Gaussian smoothing result (Color figure online).



**Fig. 6.** Shape context matching. The first and second rows correspond to reference and sensed objects respectively. The red points on the contours indicate the matched point and black points indicate the unmatched point. The third row represents correspondences between best pairs, wherein costs are minimal for two shapes (Color figure online).

Control point candidates are selected from best point pairs of matched objects. Control points are uniformly distributed in contours of objects (see Fig. 7). Affine transformation can be used as the matching transformation function for satellite images when the demand of precision is not very high. To estimate transformation parameters, the number of control points should be greater than the number of parameters to be determined (i.e., three points at least). The resulting over-determined system of linear



**Fig. 7.** Control points (red circle) and object contours overlay on the original image respectively (Color figure online).

**Table 1.** Precision statistics (pixel)

Check points	RMSE_X	RMSE_Y	RMSE_XY	Max_X	Max_Y	Max_XY
16	3.94	5.03	6.39	6.56	8.79	10.65

equations can be solved via the least squares method. The invalid control points can be excluded by Random Sample Consensus (RANSAC).

Geometric precision is evaluated by computing residual error of check points and the statistical results are summarized in Table 1. Geometric precision of the proposed method is not high. The main reason may be related to complex local deformation between multi-mode images, affine transformation alone would not provide a sufficient model for its geometric co-registration. Uneven space distribution of control points is another key influence factor. Moreover, check points are really difficult to select same point in urban SAR images and imprecise control points due to curve smoothing can also cause some loss of precision. As stated earlier, image matching focuses on searching same objects or scenes and achieves approximate alignment between two images. This study emphasizes the stable and reliable capability of object matching to solve correspondence problem. Of course, the matching result can be utilized as preliminary information for strict images registration.

## 5 Conclusion

This paper presents a novel object-based SAR image matching method, which can be considered as high-level image processing based on image interpretation and accords with human visual perception. Comparing with traditional feature-based methods, the main advantage of the proposed method is that object-based matching can avoid correspondence problem between features. The geometric shape and property information

of the ground objects can be used in the matching, rather than meaningless point, line and edge features. The shape context descriptor is used for shape matching between objects with local deformations and exhibits globally discriminative characterization for variform objects. The proposed method also can be used in matching between SAR and optical image, even image and map.

Our method is illustrated by areal water regions, which have distinct shape and are easy to segment. The experiment and many other studies not listed here for reasons of space, demonstrate water regions are more suited for multi-mode image matching even shapes of same object are inconsistent and have local deformation. Besides, other typical ground objects, such as farmland, vegetation, villages, and roads, are also can be used in image matching. Of course, these objects are more difficult to extract and may have greater difference. Moreover, shape context descriptor is not suitable for linear water region, e.g. river, and sequence matching algorithms should be adopted. All of these problems need our further endeavor.

**Acknowledgment.** This work is supported by the National Natural Science Foundation of China (No. 41331176, No. 41271425).

## References

1. Zitova, B., Flusser, J.: Image registration methods: a survey. *Image Vis. Comput.* **21**(11), 977–1000 (2003)
2. Li, H., Manjunath, B.S., Mitra, S.K.: A contour-based approach to multisensor image registration. *IEEE Trans. Image Process.* **4**(3), 320–334 (1995)
3. Suri, S., Schwind, P., Uhl, J., et al.: Modifications in the SIFT operator for effective SAR image matching. *Int. J. Image Data Fusion* **1**(3), 243–256 (2010)
4. Wang, Z., Zhang, J., Zhang, Y., et al.: Automatic registration of SAR and optical image based on multi-features and multi-constraints. In: 2010 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1019–1022 (2010)
5. Han, Y., Kim, Y., Yeom, J., et al.: Automatic registration of high-resolution optical and SAR images based on an integrated intensity-and feature-based approach. In: 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 6107–6110 (2012)
6. Ye, Y., Shan, J.: A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences. *ISPRS J. Photogrammetry Remote Sens.* **90**, 83–95 (2014)
7. Duan, X., Tian, Z., Ding, M., et al.: Registration of remote-sensing images using robust weighted kernel principal component analysis. *AEU-Int. J. Electron. Commun.* **67**(1), 20–28 (2013)
8. Yao, J., Goh, K.L.: A refined algorithm for multisensor image registration based on pixel migration. *IEEE Trans. Image Process.* **15**(7), 1839–1847 (2006)
9. Hong, D., Wu, J., Singh, S.S.: Refining image retrieval based on context-driven methods. In: *Electronic Imaging 1999*, pp. 581–592. International Society for Optics and Photonics (1998)
10. Liu, Y., Zhang, D., Lu, G., et al.: A survey of content-based image retrieval with high-level semantics. *Pattern Recogn.* **40**(1), 262–282 (2007)

11. Philbin, J., Chum, O., Isard, M., et al.: Object retrieval with large vocabularies and fast spatial matching. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007, pp. 1–8 (2007)
12. Blaschke, T.: Object based image analysis for remote sensing. *ISPRS J. Photogrammetry Remote Sens.* **65**, 2–16 (2010)
13. Ton, J., Jain, A.K.: Registering Landsat images by point matching. *IEEE Trans. Geosci. Remote Sens.* **27**(5), 642–651 (1989)
14. Sheng, Y., Shah, C.A., Smith, L.C.: Automated image registration for hydrologic change detection in the lake-rich Arctic. *IEEE Geosci. Remote Sens. Lett.* **5**(3), 414–418 (2008)
15. Dare, P., Dowman, I.: An improved model for automatic feature-based registration of SAR and SPOT images. *ISPRS J. Photogrammetry Remote Sens.* **56**(1), 13–28 (2001)
16. Xiong, B., He, Z., Hu, C., et al.: A method of acquiring tie points based on closed regions in SAR images. In: 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 2121–2124 (2012)
17. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(4), 509–522 (2002)